

1 AAGCGATAGC TGAGTGCAGC GGCTGCTGAT TGTGTTCTAG GGGACGGAGT
 51 AGGGGAAGAC GTTGCTCTC CCGAACAGC CTATCTCATT CCTTTCTTTC
 101 GATTACCCGT GCGCGGAGA GTCAGGGCGG CGGCTGCGGC AGCAAGGGCG
 151 GCGGTGGCGG CGGCAGC TGCAAGTACA TGTCAGTGACA TGTCAGTGACAT GAATCCCGAA
 201 TATGATTATT TATTCAAGTT ACTTCTGATT GGCGACTCAG GGGTTGGAAA
 251 GTCTTGCCCT CTTCTTAGGT TTGCAGATGA TACATATACA GAAAGCTACA
 301 TCAGCACAAAT TGGTGTGGAT TTCAAAATAA GAACTATAGA GTTAGACGGG
 351 AAAACAATCA AGCTTCAAAT AGAGTCCTTC AATAATGTTA AACAGTGGCT
 401 GCAGGAAATA GATCGTTATG CCAGTAAAAA TGTCACAAAAA TTGTTGGTAG
 451 GGAACAAATG TGATCTGACC ACAAAAGAAAG TAGTAGACTA CACAACAGCG
 501 AAGGAATTG CTGATTCCCT TGAAATTCCG TTTTTGGAAA CCAGTGCTAA
 551 GAATGCAACG AATGTAGAAC AGTCTTTCAT GACGATGGCA GCTGAGATTAA
 601 AAAAGCGAAT GGGTCCCGGA GCAACAGCTG GTGGTGCTGA GAAGTCCAAT
 651 GTTAAATTC AGAGCACTCC AGTCAAGCAG TCAGGTGGAG GTTGCTGCTA
 701 AAATTGCTC CCTACCTTTT CTCACAGCAA TGAATTGCA ATCTGAACCC
 751 AAGTGAAAAA ACAAAATTGC CTGAATTGTA CTGTATGTAG CTGCACTACA
 801 ACAGATTCTT ACCGTCTCCA CAAAGGTCAG AGATTGTTAA TGGTCAATAC
 851 TGACTTTTTT TTTATTCCCT TGACTCAAGA CAGCTAACTT CATTTCAGA
 901 ACTGTTTTAA ACCTTTGTGT GCTGGTTTAT AAAATAATGT GTGTAATCCT
 951 TGTTGCTTTC CTGATACCAAG ACTGTTTCCC GTGGTTGGTT AGAATATATT
 1001 TTGTTTTGAT GTTATATTG GCATGTTAG ATGTCAGGTT TAGTCCTCTG
 1051 AAGATGAAAGT TCAGCCATT TGATCAAAC AGCACAAAGCA GTGTCGTCA
 1101 CTTTCCATGC ATAAAGTTTA GTGAGATGTT ATATGTAAGA TCTGATTG
 1151 TAGTTCTTC TTGAGAGTT ATAAATGGAA AGATTACACT ATCTGATTAA
 1201 TAGTTCTTC ATACTCTGCA TATAATTGTT GGCTGCAGAA TATTGTAATT
 1251 TGTTGCACAC TATGTAACAA ACAACTGAA GATATGTTA ATAAATATTG
 1301 TACTTATTGG AAGTAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
 1351 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
 1401 AAAAAA (SEQ ID NO:1)

FEATURES:

5'UTR: 1-179
 Start Codon: 180
 Stop Codon: 699
 3'UTR: 702

Homologous proteins:

Top 10 BLAST Hits

	Score	E
CRA 108000024647144 /altid=gi 12728868 /def=ref XP_002675.2 RA...	372	e-102
CRA 18000004923424 /altid=gi 4758988 /def=ref NP_004152.1 RAB1...	332	5e-90
CRA 18000004937406 /altid=gi 131787 /def=sp P05711 RB1A_RAT RAS...	328	1e-88
CRA 18000004952860 /altid=gi 131785 /def=sp P22125 RAB1_DISOM R...	320	3e-86
CRA 18000004995539 /altid=gi 103720 /def=pir D38625 GTP-bindin...	313	3e-84
CRA 18000004967528 /altid=gi 92339 /def=pir S06147 GTP-binding...	297	2e-79
CRA 18000004880958 /altid=gi 464524 /def=sp Q05974 RAB1_LYMST R...	282	9e-75
CRA 18000004908714 /altid=gi 466171 /def=sp P33723 YPT1_NEUCR G...	253	3e-66
CRA 18000005175724 /altid=gi 7497231 /def=pir T33781 hypothetical...	253	4e-66
CRA 335001098696672 /altid=gi 11558649 /def=emb CAC17833.1 (AJ...	251	2e-65

FIGURE 1, page 1 of 2

BLAST dbEST hits:

	Score	E
gi 12867866 /dataset=dbest /taxon=960...	654	0.0
gi 12097820 /dataset=dbest /taxon=96...	654	0.0
gi 12793758 /dataset=dbest /taxon=960...	624	e-177
gi 12338056 /dataset=dbest /taxon=96...	622	e-176
gi 11977068 /dataset=dbest /taxon=96...	609	e-172
gi 10339840 /dataset=dbest /taxon=960...	517	e-145
gi 10349761 /dataset=dbest /taxon=960...	436	e-120
gi 10997958 /dataset=dbest /taxon=96...	385	e-105
gi 10996533 /dataset=dbest /taxon=96...	381	e-103

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi|12867866 Fetal brain
gi|12097820 Adrenal gland
gi|12793758 Brain neoroblastoma cell line
gi|12338056 Adrenal gland
gi|11977068 Skin melanotic melanoma
gi|10339840 Uterus leiomyosarcoma
gi|10349761 Skin melanotic melanoma
gi|10997958 Placenta
gi|10996533 Placenta

From tissue screening panels:

Whole brain

FIGURE 1, page 2 of 2

1 MSSMNPEYDY LFKLLLIGDS GVGKSCLLR FADDTYTESY ISTIGVDFKI
51 RTIELDGKTI KLQIESFNNV KOWLQEIDRY ASENVNKLLV GNKCDLTTKK
101 VVDYTTAKEF ADSLGIPFLE TSAKNATNVE QSFMTMAEI KKRMRGPATA
151 GGAEKSNVKI QSTPVKQSGG GCC (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

125-128 NATN

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 5

1	59-61	TIK
2	97-99	TTK
3	98-100	TKK
4	106-108	TAK
5	122-124	SAK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 3

1	35-38	TYTE
2	106-109	TAKE
3	127-130	TNVE

[4] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

30-36 RFADDTY

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 3

1	21-26	GVGKSC
2	147-152	GATAGG
3	152-157	GAEKSN

[6] PDOC00017 PS00017 ATP_GTP_A
ATP/GTP-binding site motif A (P-loop)

18-25 GDSGVGKS

[7] PDOC00579 PS00675 SIGMA54_INTERACT_1
Sigma-54 interaction domain ATP-binding region A signature

14-27 LLLIGDSVGKSCL

FIGURE 2, page 1 of 2

BLAST Alignment to Top Hit:

>CRA|108000024647144 /altid=gi|12728868 /def=ref|XP_002675.2| RAB1,
member RAS oncogene family [Homo sapiens] /org=Homo
sapiens /taxon=9606 /dataset=nraa /length=222
Length = 222

Score = 372 bits (944), Expect = e-102
Identities = 190/222 (85%), Positives = 190/222 (85%), Gaps = 32/222 (14%)
Frame = +3

Query: 129 GGC GSKGGGGGGGSCSDMSSMNPEYDYLFKLLLIGDGVGKSCLLRFADDTYTESYIST 308
GGCGSKGGGGGGGSCSDMSSMNPEYDYLFKLLLIGDGVGKSCLLRFADDTYTESYIST
Sbjct: 1 GGC GSKGGGGGGGSCSDMSSMNPEYDYLFKLLLIGDGVGKSCLLRFADDTYTESYIST 60

Query: 309 IGVDFKIRTIELDGKTIKLQI-----ESFNNVK 392
IGVDFKIRTIELDGKTIKLQI ESFNNVK
Sbjct: 61 IGVDFKIRTIELDGKTIKLQI WDTAGQERFRTITSSYYRGAHGI VVYDVTDQESFNNVK 120

Query: 393 QWLQEIDRYASENVNKLLVGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ 572
QWLQEIDRYASENVNKLLVGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ
Sbjct: 121 QWLQEIDRYASENVNKLLVGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ 180

Query: 573 SFMTMAAEIKKRMGPAGATAGGAEKS NVKIQSTPVKQSGGGCC 698
SFMTMAAEIKKRMGPAGATAGGAEKS NVKIQSTPVKQSGGGCC

Sbjct: 181 SFMTMAAEIKKRMGPAGATAGGAEKS NVKIQSTPVKQSGGGCC 222 (SEQ ID NO:4)

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00071	Ras family	256.4	7.7e-75	2
CE00060	CE00060 rab_ras_like	170.0	3.9e-47	2
PF00634	BRCA2 repeat.	9.9	0.39	1
PF00056	lactate/malate dehydrogenase	3.9	3.4	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00056	1/1	13	29 ..	1	18 [.	3.9	3.4
CE00060	1/2	8	64 ..	20	77 ..	86.8	8.9e-23
PF00071	1/2	13	64 ..	1	52 [.	111.9	4.8e-32
PF00634	1/1	57	79 ..	13	35 .]	9.9	0.39
CE00060	2/2	65	140 ..	110	188 ..	81.2	2.9e-21
PF00071	2/2	65	173 .]	85	198 .]	142.4	4.5e-41

FIGURE 2, page 2 of 2

6
251
501
1001
1501
2001
2501
3001
3501
4001
4501
5001
5501
6001
6501
7001
7501
8001
8501
9001
9501
10001
10501
11001
11501
12001
12501
13001
13501
14001
14501
15001
15501
16001
16501
17001
17501
18001
18501
19001
19501
20001
20501
21001
21501
22001
22501
23001
23501
24001
24501
25001
25501
26001
26501
27001
27501
28001
28501
29001
29501
30001
30501
31001

1 TTTTGGGTGT GTGTGTGT GTGTGTGT GTGCCCTTAC TAGTGACTCA
51 GGTACAGTT TTCTGAGATT TTTTTCTCC CCTCAAGACA GAATCTGCT
101 CTGTCGCCCA GGCTGGAGTG CAGTGGCTC TCGGCCACT GTAGCCTCCG
151 CCTCCCGGT TCAAGCAATT TTCTGCCTC AGCCTCCGA GTAGCTGGGA
201 TTACAGGCAC GGCCACCAT GCCTGGCTAA TTTTGTATT TTTAGTAGAG
251 ACAGTGTTC ACCATGTTG CCAGGCTGGT CTTGATTCC TGACCTCGTG
301 ATCTGTCCGT TTTGGCCTCT CAAATTCTG AGATTACAGG CATGAGGCCAC
351 CGAGCCTGGC CAGTTTCTG AGTTTTTATT TGAAATCAA ATAAGCTTTT
401 TTTTTTTTT TAATGGCTT TAGAGTCCAG GGTAACGAAC ACTTTTGGT
451 GCCTTAACTC GAACCATTC CGGTATTCC GGGGTGGTGA CCGTGTTCAT
501 TTACAGAAACC AACATGTTCA TTTCAGAAAC CAAACTCGGG TAACTTTGA
551 TAAGTTCATC AACTAAGGCC CATGGCAGAA TTTGAGGGCT AAGGGGTGTA
601 ATTAGTGTAT GGGTAGAAAT AAGTGCCTC TTTCTATATT TTGGCGTTGT
651 AGGAATTAA AGTGATTCTG CAGTAAGTCT CAGGAGACAA TTTTCTTAGT
701 TCTTAGAAGT TGGAAAGATAA ACTTTGGACA ATGTATTACA CTATGCCCTT
751 TGTAATTAAA TAACTCAAGA TAATGTGTTA AAGTTAGCG GAGATTAAA
801 TTCTGAGCT GATTAAGAG AGCTGTTAAG GCCATAGGTT TTTTAAAAAT
851 GAGTTAATAT TACTCCAGA AATTGAGGCT ACTATATAGT GATGAATTGC
901 ATATTTTTAT TGCTTATTAT TTCCAGTCT TGCAAGATGG CTCAGGGTTA
951 GTAGCAACTA AAAGATAAATT CATTACATT CAACTCTGAAG GCCGGGACGA
1001 AGGTAGGAAT TGGATTTAG GCTGCTCTG GGCTGTGTCC CTCCCATCCA
1051 TGGGATGTGG AGCCATTGAA GGTTGTGGGG TCACGATGCA GGTGCTGTCT
1101 CAGAAAGATA CATCCGACTG TGTGTGCAAA TGGGCTGGGG CGGAGAAGAG
1151 AGAGAGAGGT AGAGTCCATT TGGAGACTAC TGCAATAGCC AGGCTGACGA
1201 GTTAAGAGCG GGGCACAGTA AGAATGGAA GAAATCTAAC AGAAAATGG
1251 TAGTGCAGGG GGCACAAAT GGACGATGAC CGAACCCAGG TGGGGATGGG
1301 TGAGTGCAGCA GAAGAACCGC TCCGTGCCGT CCAGCCCCCA ACCAGCGCCA
1351 CCTCTGTTC TTAGAGCGGA CGTCCTCTA CCAGCCCCCA ACCAGCGCCA
1401 CCAGGGTGGC GCAAGCTCA AGCTGGTCAG GTCAGCAACA GCCGCAACGG
1451 AGGCAGGAGC CGACACGCTC GTACCCCGC CCCCTCCCCG CCCCCGCACC
1501 CCCGGCAGTC CCTCCGGTTT GACCACCTCC CCCGGTCCCT TGCCCTCCCC
1551 GACCCCCAGC CTCCGTCCGC CGCCGGCACC ACCCTCCGCC CCTCTCCGCC
1601 CCCTCCCCCG TGGGGCGCTG ACTCGCCCGG CTGCCACGTG TCACTGATGA
1651 CATCACTAGG GCAGCTCGGC CTTAGGCAAT CCCGCAGGGG GAGTCCGAGC
1701 GAACTCTTAG CCAGCGAGTC AGAGGGGAGG GGAGCAGGGG GGGGCCAGGG
1751 GTGGGGAGGT GAGGGAGTGG GGAATGGGGC GGGGACAAAC CCTTCAGGTA
1801 CGCATGCCCG AGAGGCCGG CGCTTGGCGG GAAAGCTGAGT CCTGGCCTTG
1851 CGTCGCACTG TCTGTCTCA GTCGCGTAG CGCGCCTCGC GACTCCCTTT
1901 CCCGGCATGC CAGGGCGTGC GGCGCCCTC TGGGCCGTG AAAGGCCCT
1951 CGGTCTAAGG CTTCCCTATT TCCTGGTTCG CGGGCGCCCA TTTTGGGTGG
2001 AAGCGATAGC TGAGTGGCG CGGCTGCTGA TTGTGTCTA GGGGACGGAG
2051 TAGGGGAAGA CGTTTGTCT CCCCACAG CCTATCTCAT TCCCTTCTTT
2101 CGATTACCCG TGGCGCCGG AGTCAGGGCG GCGGCTCGG CAGCAAGGGC
2151 GGCCTGGCG GCGCGCGAG CTGCGAGTC ATGTCACCA TGAATCCCGA
2201 ATAGTGTAGT CAGGAGAGCA CGCGTGGCT CGGTCCGTGG GCCAGCTTGG
2251 GGGATCTTAA AGGGGTGAG GAGGGTTGGG GCAGAAGTCG GGGCATCGGC
2301 TGGGGTGAGG CGAGGGTGAT GGGTCAGGAG AGGCTGGCGG CGGGAGTCG
2351 GGCCCCATTG TCTGACCGG AGGGCGGCC GGCGGGGGGG GGGGTGGGGC
2401 CGGAGGGGTG AGCCGCCCG GCCTGGACCG GTCAGGTTA GAGGGCCTGA
2451 CTGCGGGCG GGTGCTGAGG AAGCTCGCG AGGGGCCCTGG GGCCTGGTGA
2501 AGGGGTATCT TCTCTCGGAG GCAGTGAATT TTGAAGGAGG ACTTGCTCT
2551 AAGGGGAGGG GATGGGTGG GAGAGCCCTT CTAGAGGGCA CTGTCAAGACC
2601 CTGCGCCCGC ACTCTGCCGA GCTGTCAGGA TCTTCGGGGT AGAAACCA
2651 TTACTTGTAA AATCTGAGC TTGTGGTGT TCTCTCTTC CATCCTCCCC
2701 GCCAGTTTC AGGTAAATATG GATGTTTTC GGGACTGCGT GGGATTGAGG
2751 GGAATGAGTA GATGGTGAGA AGCAACTGAA CATTATTAG TTCTCTTTT
2801 GAGTTGTGTC TTGGAGGAGT TGTTAAGAG CTGCCGGGT CCATTGCCCT
2851 CCTATAAAAA CCTGGCATT TGTGAGAATT TTGTTTTTT TTTTTTTAAA
2901 GAGGACACCT AAGTCATTTT GTCTCTGTG GGTCAAGGGGA AAAAAAAA
2951 ACTAAAGCCA AGAAATGTCT TTTGATACT CGCAGATTA AGGAAGCTG
3001 CTGTCAAGTT GAAAGAGAAA CGAACGGGAC CTATGATAGA TCTGTATGTA
3051 GGTGGTGGAT TACCTGCTTG GATGCTTGCA GATAGGAAAT GAGGTTCCAT
3101 GACGTGTCAT GAAAAGTTAA TGCATTCTT TTCTTGCTT ACTCAAGAAG

FIGURE 3, page 1 of 21

3151 TCACCCACAGC AGATGTGACA CACCTGGCAC CTTTCTGGG AACTGGTGT
3201 CACTTCCCTT GGGTAGAGTT TGTTGGGCTC TCCTCAATGG CCCTTTAAA
3251 ATTTCTCTA CAGTTTACAT GCATGTAAGG TAATGAATAA TTGAAAGAGA
3301 CCGAATTGGT ATTCTTTTC AGTGTCAAAG GCCTTGAGG GATGGGGAA
3351 AATCAGTATT TGTGTTAAAA GTTGAGGTTA TTGCTGGTT TGTCATTA
3401 CTGCTAGACA TTTTCCCTA AAAGGTCAC CCACAGGT AGTCAGTGT
3451 CATATGTTG TCACATGGCT CTTGCAAAAT GCTTACAAGT TTGTAATAG
3501 TGTGCTTGA AGCTGAAATC TTGTCACTA AACAGAAACC GTAGTATTTT
3551 ATTAGAATT CATGCTTCTAG AAGTTGAGGG TAGTGTCTT GTAGTGACAT
3601 TTGCTGTGTT GACAGTTAA AAAAATTTT TTTTCAAGGG CTCCAAGGAC
3651 AAAGTTGGTT TTGACAGTT GAACGGAGGT GAACTTGAGG TTCTTAATT
3701 AGTAGTTTTC TTGGTAACAA TAAAGAACAT GGATTACTG CTTTATCGAG
3751 GTTTATAGAC CTCTACTGTT CAGGAAATT TCTGAATTG CTATATATAT
3801 GTTTATAGT GTAAATAATC CTTCAGGATT AGTTGAGAAC TTTGACAAAGT
3851 TACTCAGCCT CTGAATTTTT TTCCCTTTT GTAAATAGG ATAATTGGAG
3901 TCATTATTCC TGTCAAGGGTA GTGTTGAAAT TCAATGTAT AAAAAAGAT
3951 TTGAAAAGT GTGAGCAT TCTTCAGGTG GTATGCACTA TTTTCATGAA
4001 AGGCATTCTA TTAGTACCAAG GATTTAGGAA TATAATCCTT GCGCTTAAGA
4051 AGTTAGATA TAGGCCAGGC GCGGTGGCTC ACCTCAGTAA TCCCAGCACT
4101 TTGGGAGGCC GAGGCCGGCG GATCCCGAGG TCAGGAGATC GAGACCATCC
4151 TCGGTAACAC GGTGAAACCC CGTCTCTACT AAAAATGCAA AAAAATTAGC
4201 CGGGCGTGGT GTGAGGACCC TTGAGTCCCA GCTACTCGAG AGGCTGAGGC
4251 AGGAGAATGG CCGTATCCCG GGAGGTGGAG CTTGCACTG ACCAAGATCT
4301 GGCCACTGCA CTCCAGCTG GACGACAGAG CAAGACTCCG TCTCAAAAAA
4351 AAAATTATTT ATTGTTTGTG GACGGAGTTT CAATCTTGTG GCCCAGGCTG
4401 GAGTGCATG GCGCAAATCT CCTCTCACCG CCACCTCCGC CTCTGGGTT
4451 CAAGTGATT TCCTGCCTCA GATTCCCGAG AAGTTGGGAT TACAGGCATG
4501 TGCCACCACT CCCGGCTAAT TTGTTATTTT TGGTAGAGAC GGGGTTCTC
4551 CATGTTGGTC AGGCTGGTCT CAAACTCCCG AAGTGTACCCG CCCGCTCAG
4601 CTTCCCAAAG TTGTTGGATT ACAGGCGTGA GCCACCGCGC CCGGCAGAAA
4651 TAGATTTTAT ACATGTCAAA TACCAAGTAGA TATAGCAAAT TCCAGATGTG
4701 TGGCATGGAT GAGAGCAACA AGATTTCAAG GGGATGGTT GTTGTTGGT
4751 GCTATCTGGG TTTTGGAAAGA CTTTATAGAA GAGAGACCTG AAAGGGATT
4801 ATCAGCAATT AGATTTGGAG GAACAGAGGG AGTACTAGG AATTTCAG
4851 GGGGAGAAGA AGGAGGAATG GCTCTAAAT GACAAGGACA GTAATAAGTA
4901 AATACGGGTG CAAATCATCC TTCTTTTGA AGACTAATGA CCTCAAAGGG
4951 ATCAAACCCA GAAACAGTTT TTATATTTT TCTGGGATCA AATACATGGG
5001 TATCTGGCCT ACTATATTTG TATTCTAGAC TGTAGTAA AATAATACAG
5051 GAATTGAGA AAACCTTGC AAAAGTGTAA GTGAAAATTA CTTAGGGTGA
5101 GAGGAAGTGA GGGATATTAA ATTAGGGGAG GTCAACAAGGG CACTGAGCAA
5151 TCAGATTTTT AGTATCTGA CTTAACAGCT TTCTTTTGT TTAATGAAAG
5201 CTCTTATCT TTAAAAAAAG ATTAGAGAA AATTGGAAA ATAAAGGAAA
5251 GAAAGAAAAAG TTCTTTAGTG TTATATCAG CAAATACAAG CTACATTCTGTT
5301 TTTAACATCT TTGTTCAAAC TCCAAAGTCT TGCTTCTCT TCAATTAAA
5351 CTTTAATGGG TGGATGCTT TCCTGCTTCC AGTATGTTT CTTAATAACT
5401 AACAAATGGTA TATTAGCTAA TGTTTACAAA TGTACTCCAG ATGTTCTTA
5451 AGTTACTTTG GTTATCATT ACCAATTATAT ATTGTTCTT TTAGAAATT
5501 ATAATCTTTG TTAATGGGTT CTGCTAAATT TGTTAGTGA AATGGGATCT
5551 TGAGAAAAAA GATTCTGAAG CAACAGAAATT TTAGATTAA TATTGGTTA
5601 CATAAGAGGT GGTAGCTGTA TTACTTTTTT TGTTTGTGTT GTTTTTTTT
5651 TGAGACGGAA TCTTGCTCTG TCGCCAGGC CTTGGCCTCC CAAAGTGTG
5701 GGATTACAGG CTGAGGCCA TGTCCTGGC TGTTTGTGTT TTTTTTTGTT
5751 TTGTTTTCT TTCTTTTTC TTGTTTTCGA GTAGGAGTCT CACTCTGTCA
5801 CCCAGGCTGG AGTGCAGTGG CGCGATCTG GCTCACTGCA ATCTCTGCCT
5851 CCTGGGTTCA AGCGATTTC CTGCTTGGT CTCTGAGTA GCTGGGATTA
5901 CAGGCATTG CCACCATAAC CAGCTAATT TGTTAGAG TACCCAGCCA
5951 TCTCTAATGT TGATCAGGCT GAAGCAGGTG GATCACCTAA GGTCAAGGAGT
6001 TCAAGACCAAG CCTGCCAAT ATGGCAAAAC CCTATCTCTA CTAATACAGA
6051 AAATTATCTG GGTGTTGG CTGGCGCTG TAATCCAGC TACTCTGGGAG
6101 GCTGAGGCGAG GACAATCTC TGAACCTCGG AGTTGGAGGT TGCACTGAGC
6151 CGAGATCACA CCATGCACT CGAGCCTGGG CAACAGAGCA AGACTTGTCT
6201 CAAAAAAAAGA AAAAAAAAAGGC AATGGGTTAAAGTGAAGT GATTAATCTGAA
6251 CAGTTAAAAA AGTAGATAGA AAGGGTTAAAGTGAAGT GCTTTTTTTT GAGGATCTGA

FIGURE 3, page 2 of 21

6301 AGAAAAATGT GGATTTTT TGAGCTACGT TTTGAAGCAG GCAGTGATTA
 6351 TTTCAGCACA TTAAGAAATG CTTAACATGG CCAGGCCAG TGGCTCACGC
 6401 CTGTAATTCT CAGCACTTGG GAGGCCAG GTGGCCGGAT CATTGAGGT
 6451 CATGACCAGC CTGGCCAACA TGATGAGACA CTGCCTCTAC TAAAAATACA
 6501 AAAATTAGCT GGGTGTGGTG GTGCACGCC GTAATTCCAG CTACTCAGGA
 6551 ACCTGAGGCA GGAGAGTCAC TTGAACCTGG GAGGCCGGAG CTGCAGTGAG
 6601 TCCAGATCAT GCCACTGCAC TCCAGCTGA GGGACAGAGT GAGACTCCTC
 6651 AAAAAAAAAG AAAAAAAAAG AAAGAAATAC TTAACATTAT TCTCTGATT
 6701 ATTCTCATAA CATTTCAT AATCCACTGG CTTCCAGTGG ATTTTTTAG
 6751 TGTCAAGAAA ATAATTTGA TTGGTTCATC TTTAAGGAAT GTGTTAAGAA
 6801 TAAAGCATGT CTACCTGTCT TCAGTATACC AGCTAACTAT AGTAGGAAGA
 6851 AATATAGTAG TCTACTTAA TCAACTATAA TTCTTTAATG CAGAAAAGT
 6901 TAAAGTATTAC CCTTCTATT TTAGCCCCA TCCCCTTAAG TATATCATGG
 6951 CTCCAGAAC TCTGAAAATG TTATCAGTCT TTCAGACTTT GCTCTCTTT
 7001 CATTTTAC TCAAGAAACA TTGGACCTTT TTTTTTTTT TTTTCTTGC
 7051 ATTGTGTTTC AAAATAATTAA TAACAAAATC TAAGTGTGG AAAGTAAAG
 7101 CAGGGTGTCT TTGTAACCTT TTGGTGTGGT TTGAAAAACT CAGAAAAGTT
 7151 TAAAGAAGAA AGATAACTAG TATTCTCAT TTCCAGAATA TGATTTTTA
 7201 AATGTCTATA GAATATCACC ATCTGTAATT CTTCCGGTAA TTTAAGTATT
 7251 CAGTAGTTGT ATAAAACCTT TAAAATATAT ATATTGAGAA TTTTGTGTGA
 7301 ATGAGATGAT GAGATAATCT TGTAGGATCA TTAAAGATA AGAACTGAGG
 7351 CCTGGCACAG TGGCTCATGC CTATAATCAC AGCACTTGG GAGGCCAGG
 7401 CGGTAGATCA CCTGAGGTCA GGAGTTGAG ACCAGCTGG CCAACATGGC
 7451 AAAACCCCTGT CTCTACTAAAG CATAGAAAAA TTAATTGGGT GTGGTGTGC
 7501 CTGCGTGTAG TCCCAGCTGC TTGGGAAGCT GAGGCCGGAG AATCTTGTGA
 7551 ACCCTGGAGG TGGGCATTGC AGTGAGCTGA GATTGCGCCA CTGCACTCCA
 7601 GCCTGGCGA CAGAGCAAGA CTCTGTCCTA AAATAAGTA AAATAAAATG
 7651 AAGATAACAA CTGAAAATTC ACATTAACAA TTTTTTTGTA GCGACTGTGC
 7701 CTCCTATGTT GTGCAGGCTG GTCTCAAACCT CCTGGCTCTA AGCGATCCTT
 7751 CCAAAGCAGC GGGTGGCCA CCATGTCAG CCTGAAATTG TGCAATTAAA
 7801 AATTCTCCGC TTGGTGTGG GCGAGGGTGT TCACGCTGT AATAGCAGTT
 7851 TGGGAGGCCG AGGCAGGCAG ATCACTGAG GTCACTGTTA GACCCGCTG
 7901 GCCAATGTGG TGAACACCTG CCTCTACTAA AAACACCAAA TTAGCTAGGC
 7951 GTGGTGGTGT GCGCTTGTAG TCCCAAGCTA CTGAGGAGGC TGAGACAAGA
 8001 GAATCGTTG AATCTGGAA AAAGAGGTTG CCGTGAGGCC AGATTGGCCA
 8051 CTGCACTCCA GCCTGGGTGA CAGAGTGAGA TTCTGTCCTA AAAAATAAA
 8101 AAATAAAAT TTCCCCCTTT AATCAAAATTAA AGTTAAAATG AGGGATGTTA
 8151 GACAGTTTAAACCAATTTA TATTTAGTTAGTTAGTTTTTTTTTTTAAACGT
 8201 TGTCTTAAAG ATGGAAGTGC TTCAAAATCA AATCTTCCTT GCCAGTTCTC
 8251 TACTGGCTT CTTTTTTTTT CTTTTTGAGA TAGAGTCTCA CTTTGTCACT
 8301 GGAGTGCCTT GGCCTGATCT CGGCTCACTG CAACCTCCGC CTTCCAGGTT
 8351 TAAGTGATTC TTCCACCTCA GCCTCTCAAG TAGCTGGGAG TACAGGTGTG
 8401 TGCCACCACCA CCCGGCTAAT TTTTGATGTT TTAGTAGAGA CAGGGTTCA
 8451 CTATGTTGGC CAGGCTGGCC TCAAACCTCT GACCTCGTGA TCCACCCACC
 8501 TCAGCCAAT TGTCTGGATT ACTTGTCGTGA GCCACGCC TGGCTTCTAC
 8551 TTGGCTTTA AAGGGAAATT TGCTTCTGTGA GTAATTCTAT TTCTCAGGTA
 8601 TCTGGTCTT TTAAATTCTG GAAGCAATCT TAATAATTAA TGTATGTGCC
 8651 CTGTAATCCC AGCACCTTGG GAGGCCAGGG TGGCGAATC ACGAGGTCA
 8701 GAGATCGAGA CCATCCTGGC TAACACGGTG AAACCCCCATC TACTAAAAT
 8751 ACAAAAAATT AGCTGGCGT GGTGGCAGGC GCCTGTAGTC CCAGCTACTT
 8801 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 8851 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 8901 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 8951 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 9001 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 9051 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 9101 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 9151 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 9201 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 9251 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 9301 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 9351 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN
 9401 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN

FIGURE 3, page 3 of 21

9451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9851 NNNNNNNNNN NNCCAGGCTG GAGTGCAGTG GCACAATCTT GGCTTACTGC
 9901 AACCTCTGTC TCCCGGGTTC CAGCATTCT TCTGCCCTAG CCTCCTGAGT
 9951 AACTGGGACT ACAGGCCCTC ACCACACCG CCAGCTAATT TTATATTAG
 10001 TAGAGATGGG GTTTCACCAT GTTGGCCAGG CTGGCTCCA ACTCCTGACC
 10051 TCAGGTGATC CGCCTGCCTT GGTCTCCAA AGTGCCTAGA TTACAGGCCTG
 10101 GAGCCACTAC GTTGGCTGC TTATCAGCTT TTTACCACTT TGTCGCCACT
 10151 ACATTTGGA ATTTCCCTT GAGAATTAGG CAAAATGCC AGACTCCCCC
 10201 CGGGCCCCCG CTTTAGAGGG AGAGGGGAGC AATTAGACTA TTCCCTTGTT
 10251 TCCCTATAGA AGGTGGGCT GAGATTACTG CTTTGATATC TGGAATGTA
 10301 TTTAGGGAAG AAAATTAGG TCTTGGCCTT TCTTGGAAC CACCTGGGA
 10351 GTGTTGCAGA TTATTAATAG GGTAAATGGT GAATGATATT CAGGGAAAA
 10401 ATGGTCTGAG GGAGCCAGAG AACTAAGTGT TAGTTGTTG GCTGACTGAA
 10451 ACATGTGAGA GATAGGGTAC AGAAGAAGTA GGAAATAGTT TTCCCTGGTA
 10501 CTTCTGTGAC AGGTGGCTC AATTGGCTGG AACACCCCTAC ACTGCTTTAT
 10551 TAAATCCAAG GTTGTGATAG GTTCCAGTTA AGTTACTGT GTTCTATGCT
 10601 TGTAGATTTC CTAATTAGGA CAAGTAGTGT TAAATATGCA TGCCCTTATT
 10651 CACAAGAGGG ACCATTCTT TGGAAACATC ACTTTTAAT AATACTAGT
 10701 GCTATTTAGC ACTTACTCGG TGCCAGGCC GTGGCTATGG TTTTTTTTTT
 10751 TTTTTTTTTT CGAGACATGA TCTAGCTCTG TCTCCCAGGC TGGAGTGGT
 10801 GTAGCACAGT CATGGCTCAC TGCAGTCTCA ACCTCTGTA CTCTAGTGAT
 10851 CCTCCTGTCT CAGCCTCTG AGTAACCTGG ACCATGCCG GCTAATT
 10901 TTTAAGAGAT GAGATGTCG TATGTTGCCT ATGCTGGTCT CGAACACCTG
 10951 GGCTCAAGTG ATCCCTCCCG CCGAGCCCTC TCAAAGTGTG GGGATTACAG
 11001 GTGTGACCCA CCTCACTGG CCATCTATGG TCTTACATA GGGCATT
 11051 TGCACTGTC ATCTCAAAC ACTGATCTTC AACAGTGAACT CTCAGTGAAT
 11101 TATGTAATTTC ATGTTTCCA AGAACATGA TGGATTTAAT TTCTCTGAAT
 11151 GTATTTCCCT TGTATAATAA TAGTACTTAA GTGGAATTAC TCTTTGTCCT
 11201 TTCTACTCTC CTTATAGATA TTTCTGGTA TCTTGATTTG GGACTGTTAC
 11251 ATTTAACCCA TTTATGGTCG TGTAGCCATA CTCACGTTAC ATTTGATGCA
 11301 TCTGCTCCCT TTGTTGTCAT ATACTCATAT AACATTTGC ATAAAGTTAT
 11351 AGGCAGTTCA CACCAAGGCT GTTCATGAAC CTCAGATTAA GAATACTTGA
 11401 TTTAGGAGAT TGAAACAGA AAAGAGAATG TTAACTATCA TTATCAATAT
 11451 TAAAATGTGA AAATCTGAGA GTGACAAAGC TTAGCTTAA ATCTGGTATC
 11501 CCAAACCAT TTGAGTTTT TTTTTTTTT TTTTTTTTT GAGACAAGGT
 11551 GTCGCTTTGT CCCCCAGGT GGAGTGTAGT GGTTGATCT TGGCTCACTG
 11601 CAACCTCCAC CTCCAGGTT CAAGTATTCA TCTGCCCTCA GCCTCTGAAG
 11651 TTGCTGGGT CAAGGCTGC GCCACCCACGC CCAGCTAATT TTTGTATT
 11701 ATAGTAAAGA CGGAGTTCA CCTTATTGGC CAGGCTGGTC TCAAACCTCCT
 11751 GATCTTGTA TCCCTCCGCC TCGGCCCTCC AAAGTGTCTGG GATTACAGGT
 11801 GTGAGCCACT GTTCCCGGCC TAATTGAGT TTTAAAATGT GGAGTTAAG
 11851 ATGTTAGTCT TAAAGTGGGT TAGATGAAAT TTATAAAAAT AGTCAAATAG
 11901 CTAAATTAT AAAAGGCCAT TTGAAACAAT TTGTTGAAAT ATATAATGTG
 11951 GATAATTATG TAGTGCTTA TGTGAGAT GGTGGTTAGC ATCTGCCCTGA
 12001 TGAAGAGCAG TTGGATTTCT TACTTACTAA AGCTAGTGAAT ATCTGAAC
 12051 CAAATTAGGC ATCTTCACCA GGCTTTTG AGCCGAGCTA ACTTACTCTC
 12101 TTTTTTATT TTATTTTTA ATTAATTAAT TTTTTTTTT TTTTTTTTT
 12151 TTTGGTAGAG ACAGGATCTC CCCATGTTAC CCAGGCTTGT CTCTGGCTCC
 12201 TTGGCTCAAG CAGTCCCT ACCTAGCCT CCCAAAGTGC TAGGATTACA
 12251 GCTGTGAGCC ACTGCCAG GCTGAGCTA TTCTCTACTA ACACAAGTGT
 12301 TCTAATTAA TTTAAAGCAGT GAATCACACT TTCTTTGTA TTGGTCAGG
 12351 TTCTGGGTGC TAGTTTATAT ATGATTGAT TCATTCTGAT AGGGTTTTT
 12401 TGTTTTTTTGTT TTGTTGTTT TTTTGAGACA GAGTCTAGCT
 12451 CTGTCGCCCA GGCTGGAGTG TGTTGGCTCG ATTTCGGGTC ATTGCAACTT
 12501 CTGCCCTCCA CCCAGGCTGG AGTGCAGTGG CTCGATTTCG GGTATTGCA
 12551 ACCTCTGCCCT CCCAGGTTCA AGCGATTCTC CTGCCCTCAGC CTCCTGAGTA

FIGURE 3, page 4 of 21

12601 GCTGGGATTA CAAGCACCCA CCACCATGCC CGGCTAATTG TGTGTTTTT
12651 TAGTAGAGAC TGGGTTTCAC CATGTTGACC ACGCTGGTCT CGAACTCCTG
12701 ACCTCAAGTG ATCTGCCTGC CTTGGCCTCC CAAAGTCTG GGATTACAGG
12751 TGTGAGGCCAT CACACCAGGC CTCAAGAACT TTTTATTTT GAGACAGGGT
12801 CTCACTCTGT CACCCAGGCT GGAGTACAGT GGTGAGATCA TGCGTTACTG
12851 CAGCCTGGAC TTCCCAAGGCT CTGGTGTACCC TCCCCATCTCA GCCCCCTGGAG
12901 TAATTTAGAA TATAGACACA CACCCATGCC TGGCAGTTTG TGTATTTTT
12951 TTCTTTTTTC TCTTTTTTG TAGAGACTGG GTTTACATCG TTGATCAGG
13001 CTGGTTTGA ACTCCTGAGC TCAAGCAATC CTCACCTTT GACCTCCCAA
13051 CGTGCTGGGA TTACAGGCAT GAGGCCACTGT ACCTGGCCTT TTCTACATTAA
13101 AAAACTTTT ATTAAAAAAAC CCAAATCTTC CTTGTTGGTT TATATACATA
13151 TATACATAGG TACACACATG GAGAATTTTA CCTTGGAGGA AGGCTTGGTA
13201 AAGAAAATAG CCCTTGGGC CGGGTGCAGGG GGCTGACGCC TGAGTCTCA
13251 GCACTTGGG AGGTGAGGT GGGCGGATTG CTCAGCTCA GGAGTTCAG
13301 ACCAGCTGG CCAACACAGT GAAACCTGT CTCTACTAA ATACAAAAAA
13351 TCAGCTGGT GTGGCAGCAT GTGCGTGTAG TCCCCAGCTAC TTGGGAGCCT
13401 GAGGCAAGGAG AACTCTTGA ACCCGGGAGG CAGAGTTGC AGTACCGGA
13451 GATTGTGCTA CTGCACTTCA GCCTGCGCGA CAGACAAAAA CTCTGCTCA
13501 AAAAAACAAA CAAACAAACAA AAAAAGGAAA ATAGCCTTC TCTATCATCA
13551 GAGTATATTA AGAGTTGAGT TTTTTTTCT GTTTTTAAAT ATTGTTGTTG
13601 TTTTATTTAA ATTACAAAAC ATGGACTCTG CTTACAAATT AAGAAAATGA
13651 CTCATGTTCA AACAAAGCATA ATCAATATAA CAGTTAATAC AAGTTAATA
13701 TTGTAATATG TTTACGGAAAT AGCATGGCAA AATAGTGC AAGATTGTTGG
13751 GAAGGGCCT ATAATTTCTG TTACAGAAA GTTTTAGTTA TGTTGATTCA
13801 ACTGGAGGAG AACAGAGCTC CCAGAAGGAC TCCAGAACAC TTGATGTTG
13851 TCTGAGTGGG GTCAGCAGCA CTGAGTTCCC ACCAGCCGA AAGTTTGTGT
13901 GTGTACATTA TTTCCCTTAA CTGCCACAAAT AATCCCAGTA AGAAAATGCC
13951 CTAGTTTAC AAACAAGGAA ACAGAGGCCAG AGAAGAGTTA AATGACTTGC
14001 CCAAGGCAT TCAAAGTAAG CAACTGAATT GGAATTTTAA CTCAAAGGCT
14051 TGGATGTCCC ACTACAAACAA ATAGGCTGTT TCTGCTTAC TACATGTGCT
14101 TACTCTAAG AATTAAACAT TTTAGGCTGG TTGTGTTGGC TCACTCTGT
14151 AATCTCAGCA CTTTCGGAGG CTGAGGTGGG TAAATCATT GACCTCAGGA
14201 GTTGTGAGCC AACCTGGGG ACATGGTAAA ACCTCATCTC TACCAAAAAAA
14251 AAAAAAAAGAA CTAGCTGGAC GTGGTGGCAC GCGCTGTGG TCCCAGTAC
14301 TCAGGAGGCT GAAAGTAGGAG GATCTTTGA GCCTGGGAGG TGAGGTTGC
14351 AGTGAGCCCA CATTGCATCA CTGCACTCTA GCCTAGGTGA CAGAGTGGAGA
14401 GCCTATCTCA CACACAAAAA AAAGAATTAA AAATTTAGT CAAGTAATTAA
14451 GGCACTAACAA TTTTGTGGTC AGTTACTTTA CGAATTCTAG GTGGGAGGCC
14501 TGATGTGGTG GCTCATGCC GAAATCCCAAG CACTTGGGA GGCTGAGGCA
14551 GGAGGATTGC TTAAGGCCAA GAGTTCAAAT CAGCCTGAGC AACCTAGTAA
14601 GATCCCCCTT CTGCAAAAAA TTAAAAAAATT AGCTGGGAT GTAGTGTG
14651 ACCTGTAGTC CCAACCACTT GGAGGGCTGA GGTGGGAGGA TTGCGTGTAG
14701 CCAGGAGTTT GAGACCTGGG CAGCATATGA AGACCCCTGTC TCTAAAAAAC
14751 TAAAAATAAA AAATAGCCAG GTGTGTTGG TGTGTTGTG GTCCCAAGCTA
14801 CTCAAGAGGC TGAGGCAAGA GGTTGCTTG AGCCCAGAAG TTGGAGGCTG
14851 CCGTGAACGT TGATTGCACC ACTGCACTTC AGCCTGGGTG ACATAGCAAG
14901 ACCCTGTCTC TGTGGTGGTG GTGGGTGGGG GTGGGGGAAG GGATTTAAGA
14951 AGGGTTGTG AGGTATGTAT TATTATATAA TGGGCTTTA ACTTTACCTT
15001 TCACATCTTG GGTGAAATT AATTGTATCC ATTCTCAGTT TTCTGTCTT
15051 GCTATATATT TAAACTTGGG GACTTAGAGG TCATGGATGT CTTCTATGA
15101 AAAGCAAATG AAGCAGAGGG CTGCCTCTC TTGCTGTAGA GGGCACACTT
15151 GCTGCGAGC ATGTTACTGT TTTATGCTT GCTAGGCTT GGGAGTTGTG
15201 ACTTGTATGA TCATGACT TACAACATT AGTTGCAAT TTTTAAACTT
15251 TAACTTTAGA TTATATATGT AAACCTCTGT GTTCTTTGT CACTGATAAT
15301 CTGAACAGAA GCCTGGATA AATAATTTCG AAGTTTTGT CTGAACCTCT
15351 GAAATTGTA TTGTTATCTC ATGGTTTGC TGGGAGGAAG GAGAAATAAC
15401 AATGCCACT TACTGTGCTT CTGTATGTG CAGACAGTAT GTCTAGATG
15451 TTTCAGAACG GTGATTTGTA ATCCTGACAA GAACCTAAT TGGGTGGTAG
15501 TGGGTGCTAA TTGACCTAGA TAGATGAGGA AATTGAGGCT CATGGTGTG
15551 AGTGAATAAC TTGCAACCAAG ATCCTATGGC TGGTATGCAG TAGACCTCA
15601 ATTCAAGTAC GGGCTTCCA GGTCCAAACCC ATCGAGGCTT TTGAGGAGTAA
15651 AGGAGGTAGA GAACTGTGAC ACCCCCTTCT TGGTGTGTT TTGAGGAGTAA
15701 ACTTGTATGC ATATTAAAGA CTGTCTACCC TTTTGTCTAC TTGTTGTACT

FIGURE 3, page 5 of 21

15751 TGCTGCTTCC TTTGGTACTA CCCAAATTC TTTCAGCATT TCAGCTTGA
 15801 ATTTTTATTT TTATTTTATT TAATTTATTT ATTTTTTGA GATGGAGTCT
 15851 CACTCTGTT TCCAGGCTGG AGTGCAGTGG CGTGATATCA GCTCACTGCA
 15901 ACCTCTGCCT CACAGGTCA AGCAATTCTT CCTGCCTCAG CCTCCTTAGT
 15951 AGCTGGACT GGAGGTGCC ACCACCACGC CCAACTAATT TTTGTATTT
 16001 TAGTAGAGAT AGGGTTTAC CTTGTTGGCC AGGCTGGTT TGAACCTTG
 16051 GCCTCAAGTG ATCCACCCAC CTCGGCCTCC CAAAATGCTG GGATTACAGG
 16101 CATGAGCCAC TGCACCTGGC CAGCTTGAA TTTTTAGAAT ACTGTTCTAA
 16151 ACAGAACTAT ATTGGAACCT GGAAAATTAA TCTATTGCT CAAATACCA
 16201 AAGAAAAACA TGTAAATTAA GTGGTTGATT ATGGGAACAA TTTTTTTAA
 16251 GATGGTTCAT CTGAATGGG AGCATTTTT TTTAAATTGC TTGACTATTT
 16301 CTTAAATTG GGAGAAAAGA CCATTGCCCT CTCAGATTTC TGTTAATTGG
 16351 TCACATTGAT CATTATATT GACTGACAGG CTGCTTGTG CACAGCTGAA
 16401 GGATTGTTA ATTTTTTTA ATTAAAGA GTAATATGTG CTCACTGTA
 16451 AATTACAGT ACAGAACAT ATGAACTAAC TAAAAGTCT TACCTCTGT
 16501 CTCCAGCAAG GAGTAAGTGT TTCAACCTGA AGGTGGTT TGAATTGTG
 16551 TCTGTGGAGC GTACTTAAAG TGAGTGAAGA AGAAAAATTG ATGTCAATCA
 16601 TGATCATTG AGCTGAAGTT TTTATTGTT CACCCCTAA AGGTTATTAA
 16651 AATAGTATGT AGTTAGTAG TCTTGATAAT TTTCCCTAA GATTATTGG
 16701 CCAGTATATC AGGATTGTT TTTAAATTG ATATGTGAGC TTAGTTTAT
 16751 GCTATTTC AATAAGACAT TTAGAAGAAG ATAAAATAAC ATTCTGTCT
 16801 TAGTCTGTT TCTGCTGTA TAACAGAATA GCACAGACTG GGTAAATTAT
 16851 AACAGTAGA AGTTTATTG GCCTGTGGTT CTGGAGGCTG GGAACCTCAA
 16901 GAGCATGTT CTGCCCTTG TGCTGTGTT TCATATGGT GAAGGTGGAA
 16951 AGGCAAGTGG GTATGTCAAG ACAGAGAGCA AGAAGGGCT TGAACACTC
 17001 TTTATAACAG AGTGAETCCA GAGATAGCTA ACCCACTTT GAGAGAATGC
 17051 ATTAATCCAT TCATGAGGGC AGAGCCCTTG TGACCTAAC ACCTCTCATT
 17101 AGGCTCTGCA TCCTTAAACT GGTTTTTTT TGTTTTTTT TTTTGAGACG
 17151 GAGTCTCGT CTGTTGCCA GGCCGGACTG CGGACTGCAG TGCGCAATC
 17201 TCGGCTCACT GCAAGCTCCG CCTCCCGGGT TCACGCCATT CTCCCTGCC
 17251 AGCCTCCGA GTAGCTGGGA CTACAGGGC CCGCCACCGT GCCCGCTAA
 17301 TTTTTGTTA TTTTTAGTA GAGACGGGT TTCACCTTGT TAGCCAGGAT
 17351 GGTCTCGATC TCCTGACCTC ATGATCCACC CGCCTCGGCC TCCCAAAGTG
 17401 CTGGGATTAC AGGCGTGAGC CACCGCGCC GGCCCCCTT AACTGTTG
 17451 ATTGGGGATT AAGTATCTAA CACAGGAAC TTGGAGGATA CATTAAACC
 17501 ATAAGAATTG CTGTCATGCA AATGAATCCA TTCTAGATGA AAGAGAATGA
 17551 ATTTAGTTTC CATTGAACCTT TATAAATAGG CCTTTCTAA GGACTTACA
 17601 GCTGATATTA TAAAATTAT ATTGTTTTT ATAAATTGTT ATTGTTATT
 17651 CTGTTGTAC AAATACAATT ATACACTATA GTTCTCTGCT GTTAGATT
 17701 TTTTCTCTT TAGCATGTT CCAAAGGGT GAATGTTGAA AGTTGGTTA
 17751 ATGTCATCA GCTTCTTT GTAAAGTGT CATTGACATG TGAACCTTGT
 17801 CTGAGAATCT AAATTATTTC TCATGAAAGA AGAAAACAGT ATATTCTCAT
 17851 TTAACCCAGA ATTTAACTTC ATATACTTGT GGCTGTATTG GGAGTATGCC
 17901 ATTGCTGTCT GTTACAACC TGACCTACTC TACCTACTA GAAGTAATT
 17951 GTGTTATGAT AGGTGTGCTG TGCTGACATA TGCTGAACAT ATTGTAAGG
 18001 GTGTTAAGTC ATTGAAATAAA ACGCTTTCT CCTCCTTCA AATAACATT
 18051 TTTATTCTG GTTAAAGAAG TCATACAAGC TTACTGCAGG TTGTTAAAAA
 18101 GGTATAAGA AGAAACCGTC AATCATTAT AATCTACAG TTGACTTTC
 18151 CTGCTCCAGC CTCTAGAGT GCTGAGATGA GCTAGCCATG CCCAGCCCCT
 18201 CAAAAGATT TTTAAAAAAC AAAAATGAGG TTATACTTAA AAAAATTCTA
 18251 TATTCTTTC ACATAACAGT GTTATTGAGG AGGTTTAAAGA ATTTCAGTA
 18301 GCATTTAGA TTCAGAAACA AGCTGATTCA TCCTCTACTT TGTACTTTAG
 18351 GCAAGAAAAG AATTTCACCT AAATAGAATT TTGAACGTAA AATCTGTTT
 18401 TCTAACTTTT TATTAAAGA ATATTGTTCC ATGCTTTCAC AGTAGTGTACT
 18451 TTTAAATTAA TATTGTTA TTTTATTAT TTAGAGATGG GGGTCTCACT
 18501 CTTGTTGCCT AGGCTAGAGT GAGTGCAATG GTTCTATTCC TAGCTCACTG
 18551 CAACCTGAA CTCTGGGCT CAAGTTACCC TCCTGCCCA GCCTTCTAAG
 18601 TAGCTGGAC TACAGGTGTG CACCACTGCA CCAGGCTTT TTTAAAGGCA
 18651 TAGAAAATGG TAGTGTGTC ATACAAAAAT GGCGTAGGTA CATAACATCAG
 18701 CGGACATCAA GACTATGTT AGATCATAAA TGTACATATA TGTACCGATG
 18751 CCATTGTTGC ACGCAAACAA ATAATGGAA TTGAACCTCA AACTGAAATT
 18801 TGAAACAAAGG GTTCTGGGT GGGCCCTCTT GCTGATTGTT AATTGAATGT
 18851 ATAGTTCAAT TTTTCCCCAT CTGTTAAGCA AAAGACAATT CTAATGTTAG

FIGURE 3, page 6 of 21

18901 CAAAAATCCA CATATCCTGT CATTGATCAT TTTTCCTTA ATTTCTTAA
 18951 AGAGATGGGG CTTCTCTCA TGTTGCCAG GCTGGTCTGG AACTCTGG
 19001 CTCAAATGAT CCTCCAGCCT CAGCCTCCCA AAGTGTGGA ATTAATAGC
 19051 ACAAGCTGCT GTGCCCTGGCC CTGTCATCAG TCATTTAAC TCATGCAAAC
 19101 TGAGTAAAGT AAAACTCGTC CTTACTGTAC CTTATTGCTT TTGTTTATT
 19151 GTTGGAACCT CCAATATTGC GAAAGTAGAC CAAAGTTGA CTTATAGGAA
 19201 AAACTGATAG CAAAAATAAT TTTTCTCTT TTGCTGTATT TCATGCCAC
 19251 CATCCAGTT TAAAGCCTA CTGTTAATT CTCTCAGCCT CCTCCTTCT
 19301 GTCCAGGCTT ATTCTATGCC ATTCTTACCT TAACTGTTT TAGCTTCTC
 19351 ATAGAGTGA CTTTTAAAT TAAAATAAA TATCTGCTCG TAGTATTATA
 19401 AAATTCAAGC AGTCAACAG ATTTTTAC TAATAGAAT ACTTGACCT
 19451 CAAAAGCAGC TTATTTTAC AAACCCAGCC CAATTTGTA TTAGATTAA
 19501 CTTGAGAAAA CATGAAATGT CTCTCATATT GTTTAAAAAT ATCATAAGTG
 19551 GCTGGGACAG GTGGCTTATG CCTATAATCC CAACACTTTG GGAGGCTGAG
 19601 GCAGGTGGAT CACTTGAGGT CAGGAGTTG AGACAGGCC GGNNNNNNNN
 19651 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 19701 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 19751 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 19801 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 19851 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 19901 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 19951 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNTC ACCATGTTGG CCAGGCTGGT
 20001 CTCAAACCTC TGACCTCAGG TGATCCACCT GCCTGGCCT CCCAAAGTGC
 20051 TGGGATTATA GGCTTGAGCC TCGCCTGGCC TCCTCATATA TTTTAACTC
 20101 TTATAAAAAC CTTTCTAAA ACCCTTTTA TTTGAACTA AATTAGATT
 20151 TACTGAAATT GTGAAATCAA TGTGGAGTTC TTGTATACCC TTCTTCCGC
 20201 TTTTCTTAAT AGTAACATCT TACATACATG GTACATTGT CCAAATTAAG
 20251 AAATAAACAT TGGTACAGTG TTAACTATAG ACTTAATCTG GTTTCTCTAA
 20301 TTTTCTCACT AATGTTCTT TTCTGTTCTA GGATCTAATT CAGTACCA
 20351 TATTGATTAGT AGTGTAGGCC CATGTTAGCC ACCTTCAACT TGTGACAGTT
 20401 TCTCAGTCTT TCCTTCTTT TCGTTATCTT GACAAGTTG AAGAGTGTG
 20451 ATAGGTATT TATAGAATGT CCGTCAGTT TCTGTCAGTT TGTATTGTC
 20501 TGATGTATT TTTTTTTTT TTTTGAGATG GTGTCTCCCT CTGTCGCTA
 20551 GGCTGGAGTG CAATGGCATG ATCTGGCTC AATGCAGCCT CCACCTCCGG
 20601 GGTTCAAGTG ACTGTCCTGC CTCAGTCTCC CAAGTAACG AAACACTACAGG
 20651 CATGTGCCAC CACGCCCTGC TAATTTTTTG TATTTTAGTA GAGAACAGT
 20701 TTCACCGTGT TGCCCAGCT GGTCTCGTGC TCCTGAGCTC AGGCAATCCA
 20751 CCCGATTGG CCTCCCCAAG CGCTAGGATT ACAGGTGTGA GCCACCATGC
 20801 CTGGCCAATA TTTTGAGGA TATACTTTGG TGAGGTCATG CAGATATCCT
 20851 GTTTCCTTCTT AGTTTATCG ATTAATTAG CATTATCCA GTAAATCTTC
 20901 CTTGCAAGAA TTATTTTTTC TTTTCTTTT TCCTTAATT TTTTTTTAA
 20951 GAGATGGGAT CTCACTCTGT TGCCCAAGTT GGAATGCACT AGTGAGTTCA
 21001 TAGCTCACTG CAGCCTCAAA CTCCGGGCT CAAGTGATCC TTCTGCCCTCA
 21051 GCCTCTCAAG TAGCTGGGAC TACAGGCATA GACCACACA CCCAGCTAAT
 21101 TAAAAAAAT ATTTTGTAG ATGGGGTTTG TGCTATGTG CTCAGGCTGG
 21151 TCTTGAACCT GCTGGCCTCA TGTGATCCTT CTACCTCAGC CTTACAAGTA
 21201 GGTGGGAATT ACAGGTGTGA GCCACACAC CCAGCATTGC AGCAATTATT
 21251 AATGTAGTGC TACTGGTCA TTTCTGTTT TCTCATTTCT TCAGCATGTG
 21301 TTATTGACTT GTCTCTTCCC TCCCATTAT AATCATTAT ACTGCTATGA
 21351 ATTCACTGAGT ATTATTTTG TGAGTTATAA TCTAATACGT ACTTAATTAA
 21401 TTTTGTGCCT CAAATTGTT TGGCTTGGCC ATTTTTTTT TTTTTTTTG
 21451 AGACGGTCTC GCTCTGCTGC CCAGGCTGGA GTGCAGTAGC GCCATCTCTT
 21501 CTCACTGCAA CCTCCACCTC CCGGGTTCAA GCGATTCTCC TGCCTCAGCC
 21551 TCCTGAGTAG CTGGGACTAC AGGCCTGTGC CGCCACACCC GTCTAATT
 21601 TTGTATTCTT AGTAGAGACA GGGTTTACCT ATGTTAGCCA GGATGGTCTC
 21651 GATCTCCGTGA CCTCGTGTAC TGGCCCGCTC AGCCTCCAAA AGTGTGGGA
 21701 TTACAGGTGT GAGCCACAA GCCCCGACCGG CTCCCTGTATC CTTTTAACAT
 21751 GAGGTGCTGT CATCATTTT TCCCCCTAAT ATTTTGCCCA AAAATGTTAA
 21801 TCAAGGATGG CACAAATTT CTGTAGCTGT ATCTCACAAT GAAAGAGGCC
 21851 TGATTAAGGAAAGTAAAGTGA AATGTTCTC TGATCTCTTA GCACATGCTT
 21901 TGTAAAGGC ACAGTGCTAG ATCCCTGTAT ACGTAGATGA GTAAGTCAGC
 21951 TTACCTTCCA CACCCACAGA TAGCTATGTC AAACCTAAGG GTGGAGAAAC
 22001 ACAGACCCCA AACCTCTCGA GGGTAGAAAA TATGAGGTTA TAGTAGATTA

FIGURE 3, page 7 of 21

22051 GAACTACAAA AAGCTAGAGG AAGTTCTGAA CTGGAAACAG TGGATAGGAT
 22101 TTACTAGAAT AATTACGAG GGTGACAATT GTAAATCTC ATAGGTTCT
 22151 TTTTTTCCT TTCTCTTTT TTTTTTTGATGGAGTCT CGCTCTGTTG
 22201 CCCAGGCTGG AGTGCAATGG CGCAGTCTCT CCTCACTGCA ACCTCCGCCT
 22251 CCTGGGTCCA GGTGATTCTC CTGCCTTAGC CACCCAAGTA GCTGGGATTA
 22301 CAGGCATCTG CCACCATGCT GAGCTAATT TTGTATTTTT TTTTTAGTA
 22351 GAGACGGGGT TTCACCATGT TGGTCAGGCT GGTCTGAAC TCCTGACCTC
 22401 AGGTAATCCA CCCACCTTGG CCTCCCAAAG TGCTGGGATT ACAGGTGTA
 22451 GCCACCGCGC CCAGCCAAT TTTTATTGGT TTCTAAACTA GCGTAATTAA
 22501 GTTTTTTCA CTTAAGTCAA AATTATATTA TTGTAGGATA AAAACTTAGT
 22551 GATCCAATT CATGAGGAAT GAAGAATAAA TACATTAAA GTCTTACCAT
 22601 TTGCTAAATT AGTCTTGGCT CTTTGATACCA AAATTCTGTC CTTGTGCTCT
 22651 GTAATTCTAT ATTGTATAT TTTCTATCAA CATTTTACT GTGTGGTGT
 22701 TTGTAATTAA TAAACAGTT TAAAGCAAA CTCAGAACAA TGAATTCTCA
 22751 CGAATATTCA GTATATTTCAGTTGAGAAA TAAACTACTT CTGTAAGTAGG
 22801 TAATTAAAAA TGTCCCAATG CAAGTTAACG TGTCACTGAT CACGCTATT
 22851 AGGTGTGTGTTT CTTTGATAAG GGGAGGTGGG GAAGTTGTG GTTTGATT
 22901 TATTGCCCCCT TCTCATGTGA CTGTTGTAT GTAGTAAAC AAATGGTTG
 22951 CGAGAGAAC AGTAGTCTT TGCAAAAGATT GTCTTATACA GAGCACTCAA
 23001 TTCTTCATAT TATTATAAT GGCTTAAATT TAAGCCTTAA ATTATTAGAA
 23051 ACTCATAAAT AATTTTTTA TTGTTTTTGAGATGGGAG TTTGCCCTT
 23101 ATTGTCAGG TCTGAAGTACA ATGATGTGAT CTTGACTCAC TGCAACCTCC
 23151 GCCTCTGGG TTCAAGTGTAT TCTCCTGCCT TTGCCTCCCA AGTAGCTGG
 23201 ATTACAGGCA TCGCCTACCA TGCCCTGGCTA ATTTGTATT TTAGTAAAG
 23251 ACAGGATTGC ACCATGTTGG CCAGGCTGGT CTCGAACCTCC CAACCTCAGG
 23301 TGATCCACCT GCTTCGGCCT CCCAGAGTGC TGGGATTACA GGCTCACTGA
 23351 GCCACTGTGC CCAGCCATAA TGCGTTAAAA TAAGAGTGT ATATTGTAA
 23401 AACTAAAAA AATGTAGTGG TTGAAAAAAGG TAATTAAAAA AGAATTGACT
 23451 ATTAATTCTC TGAAACCATATA ATGTAACCTG TAGTGCAATT AGGAACCTT
 23501 CATGTTCTT CTTCTTTTC TTGTTTTTTT TTGAGATGGAT GGAGTTTGC
 23551 TCTTGTGCC TAGGCTGGAG TGTGTGATGT CAGCGCACTG CAACCTCTGC
 23601 CTCCTGGGTT CAAGCAATT CTCCTGCCTCA GCCTCCCGAG TAGCTGGGAT
 23651 TACAGGCGCC TGCCACCACA CCCAGCTAAT TTGTTGATTT TAGTAGAGG
 23701 CGGGTTTCA TCGTGTGGC CTGGCTGGT TCGAACTCCT GACCTCAGGT
 23751 GATCCACTGC ACCTGGCCCG CGTTCATGTC TTGTTAAAGCT TTATGGTGC
 23801 TCTGAAATAG AGTTGTTGAT TTGTTTTTTT TTGAGAC TCCCTTTTG
 23851 CCCGTGCTGG AGTGCAGTGG TGTGATCTGA GCTCACTGCA ACCTCCACCT
 23901 CCTGAGTTCA AGCAATTCTC ATGGGTCAAGC CTCTCAAGTA GCTGAGATTA
 23951 AAGCTGCCCA CCACCATGCC TAGCTAATT TAGTATTAGT AGTAGAGATG
 24001 GGGTTTCAACCTTCA GGTACATTGA GGTTGGCTG GAACCTCTGA CCTCAGGCAT
 24051 GAGCCACTAC GCCTAGCCTG GGTTGTTGAT CTTTAAGGTG ATACTTCAGG
 24101 CAACATCTGA GGCCCAGTAC AGTCCTTTAC TTCAACTGGC TCCAGTACAG
 24151 CAAATTCCAGG GAATGTTTTT GAGTGTGTTAC TGGATGCCCT GCGTGGAGTT
 24201 CAGGGAGATT GTTACATTGA GTCCAGTTG TGTTGAACTTCTGTTA
 24251 AAAACCTCCC TACTAAGTCC CAGCTACTCA GGAGGCTGAG GCCTGAGAAT
 24301 CACTGAAACA CCTGGAGGCA GAGGTGCGAG TGAATCCAGA TCGAGCCACT
 24351 GCACTCCAGC CTGGGCGACA GAGTGAGACT GTCTAACAAAC AAAAACAAACA
 24401 CCCCCAAAAA AACCAACCTA CTATGGTAGT ATCAATGCTG TGATAGCTT
 24451 CCTTCTTCA TACAGGTAAA TTCTTAACAT ATACTCATTG TTAATGTTCA
 24501 GTGTTCAGTA TTCTTAAGAG TATTGTTGGC CAGGCACGGT GGCTCATGCC
 24551 TGTACTCCC GCACTTTGGG AGGCTGAGGT GAGCAGATTA CCTGAGGTTA
 24601 GGAGCTTGAG AACAGCCTCC AACATGATGA AACTCCCGTC TTACTAGAA
 24651 ATACAAAAAT TAGCTGGGT TGTTAGCACA TGTCGTAAT CCCAGCTACT
 24701 TCAGAGGCTG AGGCAGGAGA ATTGCTGAA CCTGGGAGGT GGAGGCTGCA
 24751 GTGACCTGAG ATTGCTTCAC TGCACTCCAG CCTGGGCAAC AGAGCGAGAC
 24801 TCTTGTCTCA AAACAAACAA AAAAAAAAG AATATTGTTT GCCAGGCATG
 24851 GTGGCTCACA CCTGTAGTCC CAGCACTTTG GGAGGCCAAG GTGGGTGGAT
 24901 CACTGAGAT CAGGAGTTGG AGACCAGCCC GACCAACATG GCTAAATCCC
 24951 GTCTCTACTA AAAGTACAAA AATTAGCTTG AGCAACAGAG CAAGACTCTG
 25001 TCTCAAAAAA AGAAAGAAGA ATATTGTTT TAATTAAGAA GGAACCTTAT
 25051 CAATAGTAGT AAAGTCAGCC AGCTGAACCTG CCAAGTACAA ATTGTTGGTA
 25101 TTAGGTATCA ATCATTATT AAGGATAATA TTCTACAAATA GCGATTTTT
 25151 TAAAAATTAA AAAATCTCAA ACTGGAAAGG ATGTCTAGTT CATTCTATGC

FIGURE 3, page 8 of 21

25201 TTCAGTCCCC TCTTCTGATT TACTTGTGTTA GAAGATTTT GTTCCTTCT
 25251 CTGACTTCA TTTTGTGCT GACTGGCACT TGGGATTTT AAAAAATTAT
 25301 TTTCCTCAT AATAATTAAA GACAATAAGT ATAACAATAA GTATAATATG
 25351 GTAATTGCT AAAACCCAA CAATGTTTA AGTAATGCAT ATCATTATGT
 25401 AACCTACGT AATAGTTGAA TATTCAACAA GATAATCGCT TATAGAAGTT
 25451 TTATATCCTC TCTTCTTGG CAGTGAATT AAAACAAAAA AAATAAGTTT
 25501 TATGTCCTGT TTACATGTAA ATAATTAA TCTAAATTGT GACGTGGTTT
 25551 TCACTTACG ATATTTTGA AAGTAATCA AAAAGGACAA AATACAAAAT
 25601 CATGTATATC TTCTACAAA ACGATATATA AATTCTAAGG TTTTGTCCCT
 25651 TTTGAAATTG CTTAAAAGAA TGATAGAAC TGGTGTCTGA GTGGGGAGG
 25701 ATCTATGAGG GATTTCTG GAGACCGTGG GTGAATAATA ATGTTGTCTT
 25751 AGTCCCATGA AGGAATCTCT GGGGATAGTT TTTGAGTTAG GCCTGGCAAT
 25801 GTTAGAGATA CATAAAGAGA GCCTTGTGTT ATCACTGGGT GCGGTGGCTC
 25851 ACACCTGTAA TTCCAGCACT TTGGGAGGCT GAGGCGGGCA GATCATGAGG
 25901 TCAGGAGATC GAGACCATCC TGGCCAACAC GGTGAAACCC GTGTCCTACTA
 25951 AAAATACAAA AATTAGCTGG GCGTGGTGGC GCATGCCAT AATCCCAGCT
 26001 ACTCGGGAGG CTGAGGCAGG AGAACACTT GAACCAGGGA GTTGGAGGTT
 26051 GCAGTGAGCC GAGATCGCAG CACTGCACTC CAGCTGGGT GACAGAGCAA
 26101 GACTCCGTCT CAAAAAAAGG AAGCTGGTT TTCAATGGTT CTGAAAATG
 26151 CTTTAATACA AGTGTAGAGT GTTAGTCAGG TTTTCACCTT GGATAAACAG
 26201 CCTGTGAATT TATCACATT CTAGTTATA ATAAGGGCTT TCAGAAGTTA
 26251 TATGAACATT GTTTGACGG GAGAATTCAA GCTGGATGCT AGAGAAGGAT
 26301 CGTGAGAAC CCTTCATGAGG AGGAGTGCTA TGAAATTATT TGATCTTGG
 26351 ATTTTTTTT TTTTTTTTT TTTTTTTTT TTTTGAGAC AGAGTTTCTG
 26401 TCTTATTGCC CAGGCTGGAG CTGGAATGCA GTGGCACGAT CTCGGCTCAC
 26451 TGCAACCTCT GCCTCTGGG TTCAAGCAAT TCTTCTGCCT CAGCCTACCA
 26501 GGTAGCTGGG ATTACAGGCA TCGCAGAACCA TGCCCAGCTA ATTTTTGTAT
 26551 TTTTAATGGA GACGGGGTTT CACCATGTTG GTCAAGGCTGG TCTTGAACCTC
 26601 CTGACCTCAA GTGAACCTGCC TGCCCTCAGCC TCCCAAAGTG TTGGGATTAC
 26651 AGGTGTGAGC CACTGCCCT GGCGTGTACT TAGAATTGAGGAGACT
 26701 AATATTCAT GGGCAAAAC AATGAAAAGT TACCTTTCTG TATTCTAATA
 26751 CTATAGAGGA GTGGGATTTA TTTAGAATGTT TTAAAGTATC TTGGGCAGTC
 26801 CAAGAGTGGC TATCACTTAT TTTCTTTTC CTTCTTTCTT TTTAAGTGG
 26851 AGTTCACTGA TGTAGAGAT CATAGGTGGC ATTGCCTACT TTTTACATAA
 26901 TTTTATCATG TTTAGTGATC TGTAGAAGG GCTGTGGCTG TTTGCAGTTT
 26951 TGGCTTAAGC CATGCATGGG CTTTATAGGA GATGTAGTCT TCACACTGAG
 27001 TTGTATTG TAGCTGTGTT TTTGTTTTG TATAGCTTAT AGCAATGCAG
 27051 TGTGCTTTT ATTAAACATCA TTTCTTTTTT CTTTTGAG TGATTATTTA
 27101 TTCAAGTTAC TTCTGATTGG CGACTCAGGG GTTGGAAAGT CTTGCCTCT
 27151 TCTTAGGTTT GCAGTAAGTT GAAATTGAAA TGTCTTACA ATTAATGTTA
 27201 CAATTAAATGC TATGTATGTT TTCTAGGTAG ATAAAATTAA ACAGTTTAT
 27251 TCAGAATAAG TTAATTCTTC CAGAATTAT ATATTTAAAG ACTCCAATA
 27301 TACATCCCCA GTGGTATCTT GGACTGTAA ATAGAAAAAT ATTGTTGCTC
 27351 TTAAAAGAAA TTCAAGTGAAG TCTGGTTATA AAGTCAGAAAT GTCTAATACT
 27401 TTGGTCAGA GTCAAACAGC AGTTCATAA TAGGCAGCAA GTTAAAGGGG
 27451 TAGTTGGTGG CCTGTGTTGA AAGCGACTTG ATGAAAATAA ATCTTTAAAT
 27501 TAAACTTTAG TAGAATAAA AGAAAAGCA GAGCCAGGTG ACGCAGTGG
 27551 TCATGCCCTGC AGTCTCAGCT ACTCAGGGTG CTGAGGGTGG AAGGATCACT
 27601 TGAGTCTAGG AGTTTGAGA CCAACCTGGA CAACATAGCA TGACTCTGTC
 27651 TCTGAAAAAA AAAGTTAATA AAAGAAAAG TAGGGTCTTG GACAAACTTC
 27701 GTTGGCCAAT GGCATAGTT TAAATGCTGA AGCTGACAGA TAAAGGACTT
 27751 TTGACTTAAC AGAATCCCCA GTGTCCTTCA TAGTCTTTAT CAACTACCTT
 27801 TAAATTAGC ATGTTTCTG GCCAGGTGGC GTGGCTCACG CCTGTAATCC
 27851 CAGCACTTT GGAGGGCGAG ACGGGGCGAT CACAAGGTCA AGAGATTGAG
 27901 ACCATCCTGG CTAACACGGT GAAAACCCGT CTCTACTAA AATACAAAAA
 27951 ATCACCTGGG TGTGGTGCCTA CACCCCTGTA GTCCCAGCTA CTCGGGAGGC
 28001 TGAGGCAGGA GAATCGCTTG AACCCAGGAG GCGGAGGTGG CAGTGAGCTG
 28051 AGATGGTGCCTC ACTGCACCTCC AGCTGGCAA CAGAGCAAGA CTGTCCTCAA
 28101 AAAAAGAAA AAAAATAAA AAAACAAATT AGCATGTTTC CTTCTAGAG
 28151 ATCATTGTTT CTCAGAGCAT GGACCAAAGA CTCCCTGGGGG TTACCAAGAC
 28201 CCTCTCAGGT AGCCCATGAG GTCAAATAT CTTAATAATA CTAAGATGTT
 28251 AGTATTGTTA AGGAATATT TACTTGGTAA TAATACTAAT ATAAAAGATG
 28301 TTTGCGTTT TCAGTGATGA CATTGGCTCT GGTACAAAAG CATGTGGTAA

FIGURE 3, page 9 of 21

28351 AAATTGCTGC TGGCTTGGTA CACATCAAGG CAGCGCTAAG CTCCAAATTG
28401 TACTCATGGT GATGGCATTG TTTACCTCTG TGCCCTCACA GGACACAAAA
28451 CAAGCCGTGC CATTTTTATT GAAGATTGTC CTTGACAAAAA CAGTTAAAAT
28501 GATTAATTGT TGAAAAATGT TGATCCATGA GTATTCTTT AAAAATATT
28551 GTGAAGAAAT GGGAAAGTCA CATAAAACAA TGTTTTTTT TTGTTTTTT
28601 TTTTTTTTTT TTTTGAGACA GATTCTGGT GTGTTGCCA GGCTAGAGTG
28651 CAGTGGCGTC TGGCTCCAG GCTCAAGCTG TTCTCCCACT TCAGCCTCCC
28701 AAGTGGCTGG GACCTCCCAA GTGGATGCGC CATCATGCCT GGCTGATTT
28751 TGTATTTTTT TGATGTACA AGGTCTCACT GTGTTGCCA GGCTGGTCTC
28801 AAACCTCTGA GCTCAAGCGA TGCATGTGCC TCAGCCCTCC AAAGTGTGG
28851 AGAAAGCACT TTTTACTGCA TACTGGCTAG TGTGTTGGTT ATTTGGAGA
28901 AAAGAAAAGC ATTTGTAGTT TTTTGAGTTG TAAGCTGAGC TAACTGCTTT
28951 ATTTTTTCT GTGGAACACC ATTTCTTTT TTTTTTTGAGA GATGGAATAT
29001 TGCTTGTGTC CCCAGGCTGG AGTGCAGTGG CACAATCTCG GCTCACTGCA
29051 ACCTCCGCTT CTCGGGTTCA AGCAATTCTT CTGCGTAGC CTCCCAAGTA
29101 GCTGGGATTA TAGGCACCTG CCACCAAGCC CAGCTAGTT TTGTTTTTT
29151 AGTAGAGATG GGGTTTCA C ATGTTGGCCA GGCTGGTCTC GAACCTCTGA
29201 CTTCGTGATC CGCTGTCTC AGCCTCCCAA AGTGCAGGGG TTACAGGGCT
29251 GAACTACTGC ACCTGGACAT TTTTTTTTT TTTTTAATT GAAAGAACAG
29301 CTAACAGACA GATTAGAACAA GAATTGGCTA TTTGACAGAT TTTCTCAGAT
29351 GAACTGTGAT AGTCATTCA AGGGAAAGTAG CTGCAAGCAT TTGTTGGCTG
29401 AAATAAAAATT TAAGTTTATC ATGGAAAATT AGAATTGAA AAAACTTAA
29451 GTTTACCACT TGACAGTATC CTAAATACAT ATGACTTTT TGATGAGTGC
29501 CGATATTAAT GAAGGTTATT TAAAAAATAT TAAATAATGT ATAATTCTTT
29551 TTATATAACA GTAAAAAATA AAACCATGAG TACTAGAATA AAACATAGGT
29601 GGCTCTTTAA TCTGGTTTG TGAAGGTATT TTTTAAATA AGAAAAAAGC
29651 AAGAAATCAG TGCTAAATTG GACTTTAAATTGTTTAT CACAGGCACA
29701 AAAATGTTAG AAAACTAATG GCAATAGCAA ATATATATAT ATGAGGATTG
29751 GTATTCTCAA CATATAAAGC ACATTGACAT ATCAACAAGA AAAGAATATT
29801 TCTCTTAATG GAAATAGTGG CAAATACATG AGCAGTCAGT TGAAAAAAGA
29851 AGTAATACAA ATTGCTGGCT GGGTGTGGGT GGGTCACGC CTGTAATCCC
29901 AGCATTAGA GGCAGGAGCT GGCAGGATCAT CTGAGGTCAG GAGTTCGAGA
29951 CCAGCCTGAC CAACATGGAG AAACCCCTGTC TCTACTAAA ATACAAAATT
30001 AGCGGATGT GGTGGCGCAT GCCTGTAATC CCAGCTACTT GGGAGGCTGA
30051 GGCAGGAGAA TTGCTTGAAAC CCAGGAGGCG GAGGTTGTGG TGAGTCAGA
30101 TCGCACCATT GCACTCCAGC CTGGGCAACA AGAGCAGAAC TCCATCTCAA
30151 AAAAAAAA AAAAAAAA AAAAGGAAGT AATAACAAATT GCCAATAAAAT
30201 ATGGAAAAAA AAAAGGCTC AACTTTATTG TGAATTAAAG GCCTTTAAGT
30251 TAAACTTAGG TGTCAATTAA TTTTTATTAA ATTGGCAAAT ATAAAAATT
30301 AGCATAATT TTAAGCAACT CTCGGTAGGT GGGAAAGAATC TAGCTGTAGC
30351 CTCAGGTGTT TGTGCCCTCAA GGAAAACCCCT CTCTGGGATG TCCATTGCTT
30401 GAACTCAAAG GTTTCAAT AATACCTGGA AACTTTTTT AAATGCTGA
30451 TCCCCATACCT CTCAAATAT TAATAGAGAC AATCGTGAGG ACTATAATAA
30501 AGAAATGTGC AATAAGCTCT GGGGGCACAG AGGGAAGAAT CTATTGGCTG
30551 AGGAGTTGAA GAAATTGTTT GGACACTCG TATTGCCCTGA GCTCAAAACT
30601 GAAGGATGAA TAAATGCCAC ATGACCTTGG GGCTGGGGAG TAAGTAGGGT
30651 TATGCAGAGA GAGATAACTG AGGCTTTGG GCAGACGAAT AGTAACGGCT
30701 CAGGCATGGG AGTAAAGGTC ATTTAGAGAT TTACAAGAAT TCAGCATTTG
30751 TTTCTTTTTC TTTTTTTTT TTGAGATGGA GTCTAGCTCT GTCATCCAGG
30801 CTGGAGTACA GTGGCATGAT CTCAGCTCAC TATAACTCCC ACCTCCCGGG
30851 TTCAAGTGT TCTCATGCC CAGCCCTCCG AGTAGCTGGT ATTACAGGGCG
30901 TGTACTACTG TGCTGGCTA ATTTTGTAT TTTTAGTGA GATGGGGTTT
30951 CACCATGTG GTCAGGCTGG TCTCCAACTG CTGAGCTCAA GTGATATGTG
31001 CACCTCTGCT CCCCAAAGTG CTGGGATTAC AGGCGTGAGC CACTGTACCC
31051 GGCCAAGAAT TCAGTATTTC TATCCAAGTA CCTGGGGGAT AGATGTGCTA
31101 CATGAATATT TATTGCCATT ATTGTTCTCT CTGCATTTTT TTTTTTTTT
31151 TTGGTTTGAG ATGGAGTCTC GCTCTGTCGC CCAGGCTGGA GTGCAGTCGT
31201 GCAATCTCGG CTCACTGCG CCTCCACCTC ATGGGTTCAA GCGATTCTCC
31251 ATCTGGTCT CTCAGTAGC TAGGTTACA GGGGTGTGCC ATCACACCCA
31301 CTAATTGTTT GTATTTTTAG TAGAGACAGG GTTTCACCAT GTGGCCAGG
31351 CTGGCTTGA ACTCCTGATC TAAAGTGAGC CTCCCACCTT GGCCTCCCAA
31401 AGTGGCTGGGA TTACATATGT GAGCCACTGC GCCTGGCCTC TATATACCT
31451 TATAGTACCT GATACTTATT AGGCACTCAA TTACAACATA ACTTTTTTTT

FIGURE 3, page 10 of 21

31501 TTTTTTTTTT TTTTGAGACA GAGACATGCC TTGTCGCCTG GGCTGGAGTG
31551 CAGTGGCACA GTCTCGGCTC ACTGCAACCT TCACCTCCCG GGTTCAAGTG
31601 ATTCTCCTTC CTCAGCCTCC CGGGTAGCTG GGATTACAGG CGCCCGCCAC
31651 CACGTCCAGC TAATTTTTTG TATTTTTAAT AGAGATGAGG TTTCACCATC
31701 TTGGCCAGGC TGATCTCAA CTCCCTGACCT TGTGATCCAC TCACCTTGGC
31751 CTCCCAAAGT GCTGGTATTAA CAGGTGTGAG CCATCATGCC CGGCCCATAT
31801 TTCTAAAAAC ATTTCTTAT AAAATGACAT TGCCATTATC AACCTGAAA
31851 ATACATTCC ATTTGGTTGT TTTCTTGCTT AGTCTTTAA TCTAGAGTTT
31901 TATACTTAT CTTTTTATT TATATATT TTATGTCATT GACTTTTGC
31951 AGAAAAGTGA GCACTGTCC TGTAGATTGT CCAATATTCT AGATTTGTCA
32001 TTTTGTTC CTTGTGATGTC CTTATGCTTA TTTGTTGTC CCTCTTCTG
32051 TAATTAGAAG ACCTAGAACT GCACTATCCT TAGAGTAGCT ACTAGCTCTA
32101 TGTAGCTATT TAAATTTAAA TTAATTAAAA TTGAAAAAAGT TTGGTGGCTC
32151 ACACCTGTAA TCCCGACACT TTGGGAGGCC AAGGTGGAG GATTGCTTGA
32201 GTGAGGAGT CCAAGGCTC AGTAAGCTAC GATTGACTC TAGCCTGGGA
32251 GACATCAAGA CCCTGTCCTC TTAAGGGGAA AAAATAATTG AAAAATCAA
32301 AAACCTAGTT TCCTGTTTC ACAAGCTGCA TAGGGCTAAT GGCTACCATA
32351 TTGGCTAGCA CAGCTTATAG AACCTTTCCA TTGTCACAGA AAGTTCTGTT
32401 TGGCAGTGC CTTCTCATTA GACCTGATTC GATTAAGGTC CATCTTGTGTT
32451 GACAGAGTAC TTCTTAGGTG GTGCTTGTG GTTCATATGA TGATAGCCTG
32501 GTCTGTTCAT TCATATATCT TTTCACGAGA AATATTTTA TTCCATTCTG
32551 AATAAAATTT CATGGCAGGT ACTTGCAAGA AGCAGTTATA ATTTTAAAGT
32601 TTAACATTAG GTTAAAAAAAT TGACAGGAA CATATATTCA CAGGTAAAC
32651 TTGTACACAA ATGTCATGG CAGCATTATT CATAATAGCC AAGAAGTGA
32701 AACACCCCAA ATCAATTATG GAATGGATAA AATGGTGTAT ATTTGTA
32751 CATGTAATAT TATTGAGCCA ATAAAATGGG CCAGGCATGG TGGCTCACAC
32801 CTGTAATCCC AGCACCTTGA GAGGCTCAGG CAGGGGGATC ACTAGAGGTC
32851 AGGAGTTGA GACCAGCCTG ACCATCATCA CGAAACCTG TCTCTACTAA
32901 ACGTACAAAAA ATTAGGCAGG CGTGGTGATG CACGCCGTGA GTCCCTACTA
32951 CTCAGGTGGC TGAGTCATGA GGATTGCTTG GACCCGGGA GACAGAGGTT
33001 GCAGTGAGCT GAGATCATGA CACTGCACTC CAGCATGGC AACAGACAA
33051 CATCCTGCCT CAAAAAAA AAAAAAAA AAAAGAAGTA CTGTTACATG
33101 GTACAACATG GATGAACCTT GAAAACATTG TGCTAAATGA AGGAAGACAG
33151 ACACAGAGGG CCACATATT TATGATTCCA TTATACGAA ATGTCACAAA
33201 TTGGCAAAATC TAAAGAGAAA TAGATTAGT GGTTGCCAGG GAGTGAAGAC
33251 GGGTTCTTTC TGGACTGAAG AAAATGTCCT GGAATTCTGTG GTTGTAGTT
33301 GCAACCTGTG GAATGTATAA GGACCACTGA ATTGTCCACT TCAAAAGGGT
33351 GACTTTATG TTATGTGCAT TATATCTAAA AAAAAATCA TAATTAGGAA
33401 GCAAGATTGA CTTCTAAGAA AAAGCGGAGT GAAATTGTTG TTTTGTGGT
33451 AATAAATGG GTGGCTGGGT CGCAAGAGTT TTGCTGATTA GTGATTAGAA
33501 AAATTATCA TAATCATTGA AAATATAAAA TATTTTTCTA TATGATGTAT
33551 GTAAAGAATT TGGCAAGAGA TGATGTTGG AAAAAATAAA GAATGGCTAT
33601 TGAGAGATC TTAAGGAAAG AAACATACAGT TAAGTACTGC TTTGTAATCA
33651 GAATATGAAG TAAGTACTGA AAGTGGATGG AGTGGCTGTT GTCAAGCATGT
33701 TATACTTTAT ACATTCATT CATAAATTG GACTGTAGAT AAAAGTAAAC
33751 TTTTTTTTA TTTACTCTTG AACAAACAGTT TTTTTTTTC CACTTAGACT
33801 TGCATCTGCT CCACGTAAACA ATACATTTAA TTGTTAATTA TTTCCCCCTT
33851 CAGGATGATA CATATACAGA AAGCTACATC AGCACAATTG GTGTGGATT
33901 CAAAATAAGA ACTATAGAGT TAGACGGGAA AACAAATCAAG CTTCAAATAG
33951 TAAGTGAATT GGCTAGTAAT TTTTTGAAA TTATTTTGAA TAAATTTGTA
34001 ATGTATTGTT ATTTGTATA TATTTACTAT GCTAACAAAAA TTGAATGTAA
34051 AATGTCTAA GATTGATGTA CTTAAGATAG AATGGTAGAA TAAGATTAC
34101 TTAGATTAAA AATAATATT TCAAGATTAC TTAAGCTCA TTGAATTTC
34151 TGTCATGAA GCAGAGAACATC TCACTGTTTA AGTCAAACTT GGTCCCTCATC
34201 TTTTTCTTT ATCAGTGGAA ATCTAAGTTC AAGTTTACCT TGTCCCTACAC
34251 TGCAAAATGTT ATAGACCATT TTTGTTGTC TTTTACTGTG CTAAGTGCAT
34301 GGAACATTAAGGAAAGGAGATT CTTCATATGT GGCTCAAGTTG
34351 AAGAGAAGTA CTTATGTAGT TCTAAGTATT TTTATTAGAT AGTGTGCACC
34401 AACTCTGTAG AAACACAGAA TTTTGTGGAA AAAAGGAACCT TAGTTTTGT
34451 AACATGTCA TTTTACTGTG CAAACACAGC AATGCTGAAA GATTTAATGA
34501 CTTGCCTACA GTTACTGGTA GAAACAGTGC ACCGAAGCTC TGTCTCAAT
34551 ATTTGTGTC TGTGTGCCAT CCTATCCCC TTATCCATCT TTACACCCCC
34601 AGCCCCCAAT TAAATATAGG CAATTATAAT AGTCAGTTG TGCCTCTTCA

FIGURE 3, page 11 of 21

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

34651 GTATGGGTCT GAGTCCTGTC AGTGTGGCA TATCTGTGGT CTTTTAAAAA
34701 ATAAATCTCT CAGTATTTT CAGAGTAGGC TATTAGCAAG AAGTAGGCTA
34751 TAAACACAGG AAACCAGTGA CTGCCCTT TCATGGAAC GATGACACAT
34801 GGAATTGGAA CGAGTCCTGC ATTAGGAGTC AGAAGACTTA GATTTGTTGT
34851 CTTGGTTCTA GTATTACCT GTTAGAGAAT CATGGGTTTG TGTCTCTGGG
34901 GAAAAGGCCG AAGTAACCCCT GAGACCCAGT TTCCCTTCTA AAATGTGTG
34951 GATGACACCT GATTACTAA TTTATAAGCT AGTTGTGAGA ACCAACTGTA
35001 ATAGCTTGT GTATGTGACA ATACGTGTGA AAGCCCTTIG TAAACTTTG
35051 GGCAGCATAT AGATACTACT TATGATATGA CATGCCAGA TAAATGGGTG
35101 TTTGATAGGT TAAGTTGCTC CCTTTTCTTA CATGACTCTG ATGAGGAAAA
35151 GAAGGTATGT TAACAAAAGA TAGGTGGCTG TGGATATTGA TATAAGTAAA
35201 CACACTTGAT GTGCAAATT AGGACTTGCA AGGATTTAGT TTTCAGAAT
35251 AGCTTGAAT ACTTCATC AGTGAACAAA TTACCCCTCA TATTTTTTCC
35301 CACGATATAA GTACAGTCTC AACCTTTAT TTGGCACCAT AAAGAGCACA
35351 TAAAGATCTA CCCAAACTG TACTTTAAAG CACTGGTATG GAATAATTGT
35401 ATTATGTGTG ATCATTGGTG TTTATAAGAT TTGGGTGTGT ATTGTGTGT
35451 GAAACATTCA TATTTGTTA CTTTCCGTG GCTGGAAGGG ATCTTATAGG
35501 ACACTGTCTT TCATCTTGT CTGTCTTCA TCTTTAATAG GAATTCTTT
35551 TCCATGCCCTG AAGGCCTCAT TTTGAACATT TTGTTTGTGTT GTTTTTTAT
35601 TTTTGAGAT ACAGTATTGC TCTGTCTCCC AGGCTGGAGT GCAGTGGGC
35651 GATTGAGCT CACTGCAACC TCCGCCTCT GGTTCAAGT GATTCTCTG
35701 CCTCAGCCTC CCTAATAGCT GGGATTACAT GTGTGTACCA CCATGCCCG
35751 ACAATTTTT TTTTTTGAG ATGGAGCCTT GCTTGTGCG CCAGGCTGGA
35801 GTGCCAGTGG TGCATCTTG GCTCGCTGCA GCCTCCGCC CCCAGGTCA
35851 AGCAGTTCTC TTGCCTCAGC CTCCGTAGTA GCTGGGATTA CAGGCGTGG
35901 CCACACACACC CTGCTAATT TTTGTATTT TAGTAGAGAC AGAGTTTCAC
35951 CATGTGGTT AGGCTGGTCT CGAACTCCCTG ACCTCGTGT CTGCCTGACT
36001 CGGCTTCCCA AAGTGCTGGG ATTACAGGCA TGAGCCACTG TGCCCAGCCT
36051 TCCGATAATT TTTGTATTT TCGTAGAGAT GGGATTTCCG CATGTTGCC
36101 AGGCTGGTCT CAACTCCTT ACCTCAAGTG ATCCACCCGT CTTGGCCTCC
36151 CAAAGTGCTG GGATTACAGG CGTGAGGCCAC CACGCCCTGGG TTTTGAAACA
36201 TTTTAAAGA GCTTACCAATT TTTCGAAAT AGCTAGTTCC ATTTCACACA
36251 TAACTCAGC TAGGCATGTT GCCTCATGCC TGTAATCCCC GCACTTTGGG
36301 AGGCCAGGGT CAGAGAGTCA CTTGAGGCCA GGAGTCAACA TAGCTCTGT
36351 GACCAGCCTG GTCAACATAG AGACTCTATC TCTACAAAAA AAAAAAAA
36401 AAAAGTAAC CAGGTGTGGT GGTCCATGCC TGTAGTCTTA GCTCCCCAGG
36451 AGACTGAGGT GGGAGGAATG TTTGAGCCCA GGACTTCAAG GCTGCACTGA
36501 GGCAAGATTG CACCATGCCA CCCCAGCTT GGGGACAGAG TGAGAGACCC
36551 TGCTCAAAA ACAAAATAAG CTTGGGCCA GTGGCTGTCC GGGCGTGTG
36601 GTTCACGCTT ATAGTCTCTG CACTTTGGGA GGCCAAGGTG GGCAGATTG
36651 CTGAGCTCAG GAGGTCTAAG ACCAGCCTGA GCAACATGCC GAAACCTCAT
36701 CTTTGCAAAA CATACAGAAA AAAACAAAAA AAACCACAAA ACCTCTAGTT
36751 GCCAGTTATT TTTTTATTT ATTCTCTAGTG ATTCTTCTTT TTTTCTTTT
36801 TCTGAGACAA AAATTCACT TTGTCTCCCT CGCTAGAGTG CAGCGGTAG
36851 CTCACACAT GATTCTTTA GAGACATGTT AATTCTTAT ATTGAGCTGA
36901 AGCCTGTTTC TTTTACTCT GTCTCTCTT ATTCCCTCCGC CTTGTAGAGC
36951 TGCCCTGAATC AGATTAATTC CTCTTTTATT GGCAAGCCTG CCCTTCAGAT
37001 TGATCTTATC ACAACCTTCA TTCTACCTCT GAAGTCTCTA TTCTTCTTG
37051 TAATGATATT TTCAGAACCT TGTGCAATT GGGTTATTCT TACATTCTT
37101 AAATGCCCTT TATTAATTT GATTCTTAA ATCAAGTATG AGATATAACA
37151 CATGAGGTAA ATCCTGTCTT GATTGGAGC CTGAATGAAT TTCTCTCTG
37201 AACTTCAAGG GCTCATGCC CTTCTTATT ATTAATCAA GACAACCATT
37251 TGTTGTTCTA GTAGCTATAT TATTTCTAGT TTGGGTCTTA AGGTTTTGA
37301 TTTGCTTGT TTTCTTTTT TCTTTTTTTT TTTTTTGAGA CGGAGTTTCG
37351 CTCTGTTGC CCAGACTGGG AGTGCATGG CGTGATCTCG GCTCACTGCA
37401 ACCTCCGCCT CCCAGGTTCA AGCGATTCTT CTGCCCTCAGC CTCCCTAGTA
37451 GCAGGGATTAA CAGGCATGTG CCACCAACGCC GGGCTAATT TGTTTTTA
37501 GTAGAGATGG GGTTCTCCA TGTGAGGTAC GCTGGCTCG AACTCCCGAC
37551 CTCAGGTGAT CGGCCCTGCC TGGCCTCCCA AAGTGCTGGG ATTACAGTCG
37601 TGAGCCACGG CGCCTGGCCG ATTTGTTGT TTTTAATTAA AATAGGGGCC
37651 TTGGCCAGGT GCAGTTGTTC ACCCCTGTAA TCCCAGTACT TTGGGAGGCT
37701 GAGGCAGGCA GATCTCTGA GTTCAGGAGT TCAAGACCAAG TATGGCAAC
37751 ATGGTGAAAC CCTGTCTCTA CCAAAACAC AAAATTCAAG CAGGCATGGT

FIGURE 3, page 12 of 21

37801 GGTGTGTCCC TGTAGTTCAA GGTACTCAGG AGGCTGAGGT GGGAGGATTG
 37851 CTTGAGCCCG GAGATGGAGG TTGCGGTGAG CCAAGATTGT GCCATTGCA
 37901 CTCTAGCTG GGCACACAGAG CGAGACCTTG TTTCAAAAAA AAAAAAGAAG
 37951 AGGGTCTCAC TTTACACTTC TGTGACTGGT GTTTTAAAAA TCTAAACACA
 38001 GGCGGGCAC GGTGGCTCAC GCCTGTAATC CCAGCACTT GGGAGGCAGA
 38051 GGCACCCAGA TCACAAGGTC AGGAGTTCGT GACCAGCTG GCCAGCATGG
 38101 TGAAGCCAT CTCTACTAAA AATACAAAAA AATTAGCTGG GCATGCTGGC
 38151 AGGTGCCGTG AATCCCCAGCT ACTTGGGAGG CTGAGACAGG GGAATCACTT
 38201 GAACCCAGGA GGCGGAGATT GCAGTGAGCC AAGATTGCGC CATTGCACTC
 38251 CAGCCTGGTG ACAGAGCGAG ACTCCGCTG AAAAAAAA AAAAAAATCT
 38301 AAACACAAGA TTTTACTTTT ATCCTATCA TTTCCCTTG CTTGGCTTCA
 38351 GTAATCCCTC AAGTTTCTA GGTCTTTCA AAATCTTGAT TCTGTTGATT
 38401 TATATTTAA TTATCTTTT CTTTCAGCTT TTCTGTTCA GGTGTGACAT
 38451 CTGGGTCTT ATCTGAGTT TATTAGATTA TAAACATTC AGCAAGATAG
 38501 GGCAGGTAAT GAGTCCAGTT GTACACCATG GAAGGCCTCT TTCTGTGATT
 38551 GTTCATTCTAT GAGGCTTTAT GAAAATGCT ACATTACACC AGGCACTTGG
 38601 AGGTACAGA GATGAATAAA ACATAGTCCA TTAGGAGGCA GACAATGGGA
 38651 GAGACAAACA TGGGAAAAAG TTACTCTGAT TATGAGGAGT AATGAGAATT
 38701 ACATATGAAG GAAAGTATTG TTAGTACTGT TAGGATTAG TGTCAGGAAA
 38751 GTTTTCAGAG TAGCAAGGAA ACATCAGAAA TTTTACTCTT TCTGCCAGGC
 38801 ATGGTGCATG TATTATCTG TTCTCACACT GCCACAAGGA ACTGACCAAA
 38851 ACTGGGTGAT TTATTAAGGAA AAAGGTTAA TTGACTCATG TTCTGCATG
 38901 GCTGAGGAGG CCTCAGGAAA CTTACTGTGG CAGAAAGGGGA AGCAGGCACG
 38951 TCTTACATGG CAGGGCGGA GAGAGTGTGA AGGAAGTGAAGGGAAAGAG
 39001 CCCCTTATGA GACCATCAGA TCTTGTGAGA ATTCAATTCA TATCACTCGA
 39051 ATGGGGAAA CCGTCGTCA AATCCAATCA CTTCTCCATA ATCCAATCAC
 39101 TTCCCTCACT GATTACAAC TGGAGATGAGA TTTGGGTGG GACACAGAGC
 39151 CAAACCATAT CAGTGCCTGT AGTCCCAGTT ACTTGGAGGC TGAGGCAGGA
 39201 GGAACACTTG AGCCCAAGGAG TTCAAGATCT GCCTGGCAA CATAACAATA
 39251 CCTCCATTG GGTAAAAAAG GAAATTTAC TTTTGGGTG CCATTGCTTA
 39301 GTTTAATCAG CTGTAACCTC TTGTTGACTT TTAGTCAAA AACAATTTT
 39351 CCTTCTATCT TTGTAAGAAGA GGTTGGTGG CAAGGAAGAA AAGGAAACTT
 39401 GCTTTATGAG GCAGCTTCTA TAGTCAGGCA CATTTCACAA ACATTAGTTC
 39451 ATTTAAACCC CTTTAGCTGT TGTACAAGGT GAATGCTATC TAGCATTTAC
 39501 AGATGAAGAA ACTGTTAGGT GACTCTCCCT AATATTAAAT AACCAGGAAC
 39551 CTGGATTGAG TGTTTGAAAG TCAGGGTAGC TTGATCCTCG AGTTCATGCT
 39601 TCCTCCAAGG ATACACTGAA AGACTTTGAG CCTCTTTTT TTTTTTCTC
 39651 TTTTTTGAG ACAGGATCTG GCTCTCTTGC CCAGAGTGC GTGGTGTGAT
 39701 CTCAGCTCAC TGCAACCTC GCCTCTGGG CTCAGCGAT TCTGCCCTCAG
 39751 CCTCTCGAGT AGCTGGGACC ACAGGGCAC GGCAGCATAC TTGGCTAATT
 39801 TTTGGATTTC TAGTAGAGAC AGGGTTTCAC CATGTTGGTC AGGCTGGTCT
 39851 CGAACTCTG AGCTCGTAAT CCGCCCGTCT CGGCCCAACA AAGTGTGGG
 39901 ATTACAGCG TGAGCCACCG ACCCAGTCCC AACAGTTTT TAAAACCCAG
 39951 AACTATAATG CAATAATGTT AGCATTGTT TTGGGAGTTT GAGCCTAAAT
 40001 GGTTGAAGTG CAGTAAATTG TTCTTAAAT ACGTTTATG AAAGTATTG
 40051 GAGTCTCTC CTTACATTTC TTCTCTTAGC ATGAAGACAA CACCTAGCCA
 40101 GGCATGGTGG CTCATGCCAG TAATGCCAGC ACTTGGGAG AATGAGTTAG
 40151 GATAATTGCT TGAGTCCAGG AATTGAGAC CAGCTGGG AATGTAGCGA
 40201 GACTCTGTCT CTACAAAAA GAAAAAATTA GCGGGTGTG GTGGCATGTG
 40251 CCTGTAAGTCC CAGCTACTCA GGAGGCTCAG GTGAAAGGAT TGCTTGAGGT
 40301 GGGAGGTTGA GGCTGCAGCG AGCCATGATC ATGCCACTGT ACTCAGCCTG
 40351 GATGACAGAA TGAGACGCTG CTTGAGAGGG GAAAAAAAAG ACACCTGCTT
 40401 GGGATGATTA AAGTTCTGTC TTGACTGGTA GTTATTGAA TTAGGTCCCT
 40451 CCAGTCCTT TAATCATGGT AGAATGTTGCT AGCAAGTGGAG TTTGCTTAC
 40501 ATGGAAGAGT TCTGTTCA AGGGCTTCG GCCAGTGGCA TTCCCTAAACA
 40551 CAGTGTAAA GGCGGTAGGG AATGTGAAAA GTATGACATA GTTCCCTGCTC
 40601 TCAACAGCTT GTAAATTTCAG TATTATTATC GTAAGCTCAA TTGTAGGTAC
 40651 TACTCTTTT CTGGACTTTC AGGTGCTTAT TACCGTGCAA TTTAGTGGTA
 40701 TGAGTTGAGG ACTAATGTTT CTATATCACA TCCCTGATAAT CTCCACAGTT
 40751 ATGAAAACAA AACTATTTC CTCCTCTCCT ACACCTTCC CCAACTTTAT
 40801 TTTAATGGAA TTGTTGGAT TTCTGATTG TTTTGTAAATA GTGGGACACA
 40851 GCAGGCCAGG AAAGATTTCG AACAACTCACC TCCAGTTATT ACAGAGGAGC
 40901 CCATGGCATC ATAGTGTGTC ATGATGTGAC AGATCAGGTA AGTTCCAAGA

FIGURE 3, page 13 of 21

40951 GGAGATTGTG TTACAGTGAC CAAGTAGGAA GCCATTATTT GATTAATGTC
 41001 AGATTCACTT ACTACTTCAT ATATAAGCCA TCAGTATTAA TTTTATGGCA
 41051 GAAAACCTTG TCCACTCTCA AATATAAATG TGAATCACCT AAAAGACATT
 41101 TGTTTCTCTG TAATAAATAA AAGATTAGTA ATTAGTTTA CGTTTGCCTT
 41151 CAAGGGATTG TGTTGTATT TATTGTCAAC TAAATAACCT TGATCAAATA
 41201 GCCAAGACTC TAACATATAG GCAAGAGTT GTAGGAAATC GTGAGTTGCT
 41251 TGGCTTATAC TGTTCTTG GTGTTAAGTA TTAACAGGAA TATGCCCTGG
 41301 TAATTAGAAC TTGTCATCA GAATTGCCAA AAGTGGATT CGGGGCTC
 41351 TGCCTATGGA GGATGTGGTT CAGAAATAAA GAATTGAAT AGGATAAGCT
 41401 GTAGGAGGAT CTTAGTATGA GAATGAGTAT CTGAAGATTA GCTGTGAGAG
 41451 AGGGCAGAGC GATGGAGGG ACAATGTGGG ACAGTGTGAA GCATGTGATC
 41501 CAGGGCCAT AACTTTTTT GTTACTATT TTTAAATCA GAAACTTAA
 41551 TTTCAGTGTCTTCTATCA AAGAAAAGGA CAAAAGATAA ACGTTCAAAA
 41601 TTGGAATTAA TTTTCTTT GGCAAATGTT AAATCTCAC TCTAATGAGA
 41651 AATCATAGCT AATTAGGAGA TAACTTACAT GTAAGCATTT AGATTCACTG
 41701 CCATTAGAAC TGCTGGTGG GTGATATCTG CAGGAGAAA AAATGATGCT
 41751 AGTTTAAAGA ATCTCTACTA TTACCGTAA ATATTTTAA ATGAAAACCTT
 41801 TCGTCCTCTA AATATGACTG TGGAAAAGAA AATGAGTATA TTTAATAACA
 41851 TCTTTGACA TCTCTAGTAG TAACAGTAGG TCATCTTATT CATAAACCAA
 41901 AATTTCACCA AATTTCAGGC CAGGCGCAGT GGCTCATGCC TGTAATCCCA
 41951 GAACTTTGGG AGGCCGAGGC GGGCGGATCA CCTGAGGTCA GGAGTTAGAG
 42001 ACTAGCCTCG CCAACATGGC AAAATCCCAT CTCTAGTAAA AATACAAAAAA
 42051 TTAGCCAGGC GTGGGGGCC GTGCTGTAA TCCTAGCCAC TTGGGAGGCT
 42101 GAGACAGGAG AATCGCTTGA ACCCAGCGGG CAGAGTTGC AGTGAGCCGA
 42151 GATGCCGCCA TTGCACTCCA GCCTGGATGA CAGAACAGA CTTGCTCTCA
 42201 AAAAAAAAGA AAAAAAAAGA AAAAAAAATTA ATCAAATTTC AAAACCAAGGT
 42251 TTGTTAGTAC ATTTAAATTG CATAATTCCAA AGCAGTTGGG TTTGCCTGCG
 42301 TTGCACTTAA ATATTAAGCT ATACTTCCCT TTCAAATAAG GTATTTCAT
 42351 CGTTAACCCCT GTAAATTCTA GTTTGTCTT GTTTAGATAT TTATAGTCAT
 42401 TTAAATTATAT CTGTTTACGG CCAGCTGCAA TGCTAACAC CTCATAACTC
 42451 AGCACTTTT GAGGCCAAGG TGGGCCGATT GAGCTCAGGA GTTCGAGACC
 42501 AGCCTGGCA ACATAGTGA ACTCCATCTA TACAAAAAAAT CCAAAAAAAAGA
 42551 AAAGACAGGT GTGGTGGCAT GTGCTGTAG TCCAGCTAT CCCGAGGCG
 42601 GAGGCGGGAG GATGGCTTGA GCTTGGGAGG TCGAGGGTGC AGTGAGCTGT
 42651 GATTGTGCCA CTGCACTCCG GCCTAGGTGA CAGAGCAAGA CCCTGCTCTCA
 42701 AAAAAAAAGA TCTCTTCACT CCTTAGCAGT GGTTATTTC TAGCTAGAGT
 42751 TGTCTCACTA GCTCTTTGTT ATTTGTCTGT TAGTCAGGA ACGATGTTTC
 42801 TGTTTATTCC AGAACTATAT TATCGAACTA TATTATCAGT CTTCAAAATG
 42851 TCTTTTCTTACGG AGTCTTCAA TAATGTTAAA CACTGGCTGC AGGAAATAGA
 42901 TCGTTATGCC AGTGAATAATG TCAACAAATT GTTGGTAGGG AACAAATGTG
 42951 ATCTGACACCA AAAGAAGATA GTAGACTACA CAACAGCGAA GGTATGTTA
 43001 AAGTTAATT TTCAACTGAA ATTGGAAGGT GTGAAATTAT GTATGGGTT
 43051 TGCAGTAACA GTAAGGCCAC AGCCTTTAA AAATATGTGC ACTAGAATAC
 43101 TGTGACAGTG ACAATTGTG TAGCATCTGT TTGGATCCAA TGAACCTAGT
 43151 TCCTCACGCT CCATTATGGA TGGTAGAAAT GCAGTAAGAA TTAGTAAAAA
 43201 AGATTTCATCA GTGTTAATTG TGCCCTCATTA TTCTCTTCTG AATTGCTGA
 43251 TTCCCTGGAA ATTCCGTTT TGGAAACCGAG TGCAGAAAT GCAACGAATG
 43301 TAGAACAGTC TTTCATGACG ATGCCAGCTG AGATTAAGAAG GCGAATGGGT
 43351 CCCGGAGCAA CAGCTGGTGG TGCTGAGAAG TCCAATGTTA AAATTCAAGAG
 43401 CACTCCAGTC AAGGAGTCAG GTGGAGGTTG CTGCTAAAAT TTGCTCCAT
 43451 CCTTTCTCA CAGCAATGAA TTGCAATCT GAACCCAAGT GAAAAAAACAA
 43501 AATTGCCCTGA ATTGTACTGT ATGTAGCTGC ACTACAACAG ATTCTTACCG
 43551 TCTCCACAAA GGTCAAGAGAT TGTAATGGT CAATACTGAC TTTTTTTTTA
 43601 TTCCCTGTAC TCAAGACAGC TAACTTCATT TTCAAGACTG TTTAAACCT
 43651 TTGTTGCTG GTTGTATAAA TAATGTTGTG AATCCTTGTGTT GCTTCTG
 43701 TACCAAGACTG TTTCCCGTGG TTGGTTAGAA TATATTGTT TTTGATGTTT
 43751 ATATTGCGAT GTTGTAGATGT CAGGTTTAGT CTTCTGAAGA TGAAGTTCAG
 43801 CCATTGCTGA TCAAAACAGCA CAAGCAGTGT CTGTCACCTT CCATGCATAA
 43851 AGTTTAGTGA GATGTTATAT GTAAAGATCTG ATTTGCTAGT TCTTCCTG
 43901 AGAGTTATAA ATGGAAAGAT TACACTATCT GATTAATAGT TTCTTCATAC
 43951 TCTGCATATA ATTGTTGGCT GCAGAAATT GTAAATTGTT GCACACTATG
 44001 TAACAAAACA ACTGAAGATA TGTTAATAA ATATTGTACT TATTGGAAGT
 44051 AATATCAAAC TGTATGGTGA TAAGTATTGT TTGATTCTT ATGGTTAAAG

FIGURE 3, page 14 of 21

44101 GGAAATAGAG CCTTGCATTA TATTCAACAC AGCCATTGT GTGTGCACAA
 44151 TGCAAACATAA GGTATTCTAG ACCTATCTTA GAGCAGCAGC CAGTATTGC
 44201 TTTCTAGATA ATATGCCAA TAACATGACC TAGAGGGCT TCTGTGCTGT
 44251 GTAGGGATTT AACCAACTTC AGTGGTTCAG GGAGCTCAA CTATATGTAA
 44301 AACAAAGTTTA GAATGTATGC TATCTAGCCC GTTATCTCTG ATCCTTCTCT
 44351 AAAACCATTT GAAATAGCTT CATTGATCAA CATTCTATAA ATGCATCTGT
 44401 GGTAGAGGTA GAAAGCAGCA CCTTTCTAA TTGGCAAATG ATCAGACTAA
 44451 TGTTGCTAA TTGTTTTCTT CCATGCTTC AGTCAGATT AACTATTTA
 44501 TCCTCCACAG TTGCTTAAC TGGTGTGGA GGAGGGTTA AGCATTAAGA
 44551 TAGGAAGCAG GAAATTTGAT TGCTCTAAAT TTAGAAATTA TATCCCTAA
 44601 AATTAAAACA TGAATACTGG GTGGTAATGA TAATTGAGGC AAATGTATT
 44651 ATTTGGTGA CATTTCGAT ATATGAAGAT TTTCTGAAAT AGGACCTTC
 44701 AGATCCTAGG GGTTTTGTT TGGTTTTAA TTGTGAGGAA TAAAAAATCT
 44751 TCTGCCACAA CTGGCATTTT AAGGTGACTG AGGTCAAACG TTGTTTCCTT
 44801 AGGTGAAAT AGCAGCCAA ACATTCTTC CGCAGGGCT TGGGATATGG
 44851 CTCTGGCAA CACATTGTG TGTGGGCTCC TTAATTAAAT GATAAAATTT
 44901 AAGCTAAACAA CAAGCCAAA ATGAATAGGT TTTTTAATT TTTATTTTC
 44951 ACTAAACAGG CAATTGAAAT ACATGGTACA AAAATAAGT GTAAGATAAT
 45001 TGTAATGAA ATAGGACAGA ATATTCAATT TTCCATCTAT GAAAATTCA
 45051 CAATAAAAT CATAGTTTAC TTTGTATTAT AGGCCTGCTT GGTGGATCTA
 45101 TTACATCTCA CATAAGGCAA CTGACAAAT CCTGAAGTT CCAATAGTTA
 45151 TTGTTGGAA GATCTTTAAT GCTTCAGAAG TTTGTTTTT GCCTTAATAC
 45201 AGTAAAGG GGGAAAGAGT TCAGAAACTA TTTCTAAAG TAGCTAAATG
 45251 ACACAAAACA ATAGTCAAGA TACTGTGATG CCATGCCGTG CACTTCAATT
 45301 TTACACAGTA AAAGTTGTTT AAATTGTCAG CTTATTCTTGT GTGAGTTAGC
 45351 GGAAACATTA CATGAACCTTA AGATGAGCAT ATTTACAGAC TTAAGTTGG
 45401 AAAATTCCAG CGTCTTTTC CCCATGGCAG TAAAGATTGG GATTACAAAC
 45451 AAATTTCAGC ATGCCCTTAAG ATTTGCTTCT ATGTATAACGC CAATAATGT
 45501 GGTCTGGAA AAAATATATA CCCCTTTATA CCCCCATTAA CAAGTACAAA
 45551 CGGTCAAAAG CTACTACAGG TTTTAATAAT CTGTTCACTT AGTAAAGGGA
 45601 ATTACCACTT GTTCTAAATA TAAGGTGCTG CCATAAAATTA GTTACATAG
 45651 TGAGGAAGAG TTGTTCTAAA TCTAAGCAGC TGACACACTCT GTGAAATCCT
 45701 TTCAGAATGA TAGTCATTGT GGTCTGAGCA GTAATTCTCT ATTCTTCGAC
 45751 CTTGGATTGA ATTTCCCTTA GCCTACATCT TGCCCTTCCA GCATATCTTA
 45801 CCTCAAACCT TCTTGTGTT CCATTCCCAC CTAAGCTCA AAATAGCCCT
 45851 GTGTTGACGT CGTCTTCCAT TTGCTGAGCT TACCTATGGA TCTCCAAGAA
 45901 CCCAGATCTT GAAACTGCTG ATCCAGCTT GAGTATCATC ACTTCCCTGT
 45951 GGATTTAACT TCCATTAAATT TTAAGGGACT ACTAAGTTAT TCCAGTGTGG
 46001 CATCACAGTG CAGTTAGCAA GCTCAGCTAC TTGACTCTAA TTGGCCATG

(SEQ ID NO:3)

FEATURES:

Start: 2181
 Exon: 2181-2203
 Intron: 2204-27090
 Exon: 27091-27163
 Intron: 27164-33853
 Exon: 33854-33949
 Intron: 33950-42859
 Exon: 42860-42991
 Intron: 42992-43239
 Exon: 43240-43434
 Stop: 43435

CHROMOSOME MAP POSITION:

Chromosome 2

ALLELIC VARIANTS (SNPs):

Position	Major	Minor	Domain	Protein		
				Position	Major	Minor
397	T	-	Beyond ORF(5')			
2326	A	G	Intron			

FIGURE 3, page 15 of 21

3486	C	A	Intron
6651	-	A	Intron
8190	T	-	Intron
8281	T	C	Intron
11546	A	G	Intron
11670	C	T	Intron
11688	A	G	Intron
14938	A	C	Intron
22261	G	A	Intron
22852	G	A	Intron
27253	A	C	Intron
28098	-	A	Intron
28597	G	T	Intron
31431	C	T G	Intron
35704	C	T	Intron
35728	C	T	Intron
36690	C	T	Intron
41002	G	C	Intron
41033	A	G	Intron
43161	C	T	Intron
43765	A	G	Beyond ORF(3')
44713	G	T	Beyond ORF(3')
44831	C	T	Beyond ORF(3')

Context:

DNA

Position

397	TGCTCTGCGCCCAGGCTGGAGTGCAGTGGCTCTCGGCCACTGTAGCCTCCGCCTCCC GGGTCAGCAAGCAATTTCCTGCCTCAGCCTCCGAGTAGCTGGGATTACAGGCACGCGCCA CCATGCTGGCTAATTTCGTTAGTTAGTAGAGACAGTGTTCACCATGTTGCCAGGC TGGCTTGAATTCTGACCTCGTGTCTGTCGGTTGGCCTCTCAAATTCTGAGGATTA CAGGCATGAGCCACCGAGCCTGCCAGTTCTGAGTTTATTGAAATCAAATAAGC [T, -] TTTTTTTTTTTAATGGGCTTAAAGTCCAGGTAACGAAACACTTTGGTGCCTATT ACTGAACCATTAGGGTATTCTGGGGTGGTACCGTGTTCATTCAGAAACCAATGT TCATTTCAGAACCAAACCTGGTAACCTTGATAAGTTCATCAACTAAGGCCATGGCA GAATTGAGGGCTAAGGGTGTAAATTAGTGTATGGTAGAAATAAGTGCCTTCTTCTAT ATTTGGCGTTGAGGAATTAAAGTGTTCAGTAAGTCTCAGGAGACAATTTCCT
2326	GCTGATTGTGTTCTAGGGGACGGAGTAGGGGAAGACGTTGCTCTCCGGAACAGCCTAT CTCATTCCTTCTTCGATTACCGTGGCGGGAGTCAGGCGGGCTGCAGGA AGGGCGCGGTGGCGCGGCCAGCTGACATGTCAGCATGAATCCGAATAGT GAGTTCAAGGAGACGCCGCTGGCTGGGTCCGGCCAGCTGGGGATCTTAAAGGG TCGAGGAGGGTTGGGCAGAACGTCGGGCATCGGGCTGGGTGAGGCAGGGTATGGTC [A, G] GGAGAGGCTGGCGCGGGAGTCGGGCCATTGTGACGCCGGGGCGGCCGCG GGGAGGGTGGCGGGAGGGTGAGCCGCCGGCTGGACCGGGTCAGGTTAGAGGGC CTGACTGGGGGGGGGTGCTGAGGAAGCCTGCCAGGGCTGGGGCGGTGTGAAGGGGT ATCTCTCTGGAGGAGCTGACTTTGAAGGAGGACTTGTCTTAAGGGGAGGGGATGG GTGGGAGACCCCTAGAGGGACTGTCAGACCCCTGCCGCACTTCGCGAGCTGTC
3486	CTGGGAACGGTGTCACTCCCTGGTAGAGTTGTTGGCTCTCTCAATGGCCCTT AAAAAATTCTCTACAGTTACATGTAAGTAATGAAAGAGACCGAA TTGGTATTCTCTTCAGTGTCAAAGGCTTGAGGGATGGGGAAAATCAGTATTGTTG TAAAGTTGACTTTATTGCTGGTTGGTCAATTACTGCTAGACATTCCCCTAAAGG TCCACCCACCAAGTTAGCTGACTGTATGTGTCACTAGCTCTGGCAAATGCTTA [C, A] AAGTTTGAAATAGTGTGGCTGAAGCTGAAATCTTGCACAAACAGAAACCGTAGTA TTTATTAGAATTCTAGCTTAAAGTGTAGGGTAGTGTCTGTAGTGTACATTGCTG TGGTACAGTTAAAAAATTCTCAAGGGCTCAAGGACAAGTGGTTTGAC AGTTGAACGGAGGTGAACCTGAGGTTCTTAATTAGTAGTTCTGGTAACAATAAGA ACATGGATTACTGTTATCGAGGTTAGACCTACTGTTAGGAAATTCTGAA

6651 TTTCAGCACATTAAGAAATCCTAACATGGCCAGGCCAGTGGCTACGCCCTGTAATTCT
 CAGCACATTGGAGGCCAGGTGGCGGATCATTGAGGTCAAGCAGCCTGGCAAC
 TGATGAGACACTGCCCTACTAAAATACAAAATTTAGCTGGGTGTTGACGCC
 GTAATTCCAGCTACTCAGGAACCTGAGGAGAGTCACTTGAACTGGGAGGCCAG
 CTGCAGTGAGTCCAGATCATGCCACTGCACCTCAGCCTGAGGACAGAGTC
 [-, A]
 AAAAAAAAAAAAAAGAAAGAAATACTAACATTATTCTCGTGAATTCTCAT
 ATTTTCATAATCCACTGGCTTCACTGGGATTTTTGTCAAGAAAATAATTGAT
 TGGTCATCTTAAAGGAATGTGTTAAGAATAAAGCATGTCACCTCTTCAGTAT
 GCTAACTATAGTAGGAAGAAATATAGTAGTCACTTAGATCAACTATAATT
 AGAAAAAGTTAAAGTACCTTACCTTACCTTACCTTACCTTAAAGTATAT
 AGGGAGGCCATGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
 8190 AGACGGGCTGGCAATGTGGTGAACCCCTGCCTCTACTAAAAACACAAATTAGCTAGG
 CGTGGTGGTGTGCGCTGTAGTCCAAGCTACTGAGGAGGCTGAGACAAGAGAATCGCTT
 GAATCTGGGAAAAGAGGTTGCCGTGAGCCAAGATTGCCACTGCACTCCAGCCTGGGTG
 ACAGAGTGGAGATTCTGTCCTAAAAAATAAAAATTCCCTTAATCAAATT
 AAGTAAAATGAGGGATGTTAGACAGTTAACCATCAAATTAGTTAGTTAGTTT
 [T, -]
 TTTTAAACGTTGCTTAAAGATGGAAGTGCCTCAAATCAAATTCTCCTTGCAGTTCTC
 TACTTGGCTCTTTTTTCTTTGAGATAGAGTCTCACTTGTCACTGGAGTGCCTT
 GGCGTATCTCGCTCACTGCAACCTCCGCCTCCAGGTTAAGTGATTCTCCACCTCA
 GCCTCTCAAGTAGCTGGAGTACAGGTGTTGCCACCACCCGCTAATTGTTAGTT
 TTAGTAGAGACAGGGTTCACTATGTTGCCAGGCTGCCCTAAACTCCTGACCTCGTGA
 8281 CTGAGGAGGCTGAGACAAGAGAAATCGCTTGAATCTGGGAAAAGAGGTTGCCGTGAGCCA
 AGATTGCCACTGCACTCCAGCCTGGGTGACAGAGTGGAGATTCTGTCCTAAAAAATAAA
 AAATAAAAATTCCCTTAATCAAATTAAAGTTAAAGGATGTTAGACAGTTT
 AACCATCAAATTATTTAGTTAGTTTAACTGTTGCTTAAAGATGGAAGTGC
 TTCAAATCAAATTCTCCTTGCAGTTCTACTGGCTCTTTTTTTGAGA
 [T, C]
 AGAGTCTCACTTGTCACTGGAGTGCCTGGGTGATCTCGCTCACTGCAACCTCCGC
 TTCCAGGTTAAGTGGATTCTCACCTCAGCTCTCAAGTAGCTGGAGTACAGGTGTT
 GCCACCACCCGCTAATTGTTAGTTAGAGAGACAGGGTTCACTATGTTGCC
 AGGCTGGCTCAAACCTCTGACCTCGTGAATCCACCCACCTCAGCCAATTGCTGGATT
 CTTGTTGAGGCCACGCCCTGGCTTACTGGCTTTAAAGGAATTGCTTCTGAG
 11546 GTTACATTTAACCCATTATGGCTGTAGCCATACTCACGTTACATTGATGCATCTGC
 TCCCTTGTCTATATACTCATATAACATTGCTAAAGTTAGGCAGTTCACACCA
 AGGCTGTTCACTGAAACCTCAGATAAGAATACTGATTAGGAGATTGAAAACAGAAAAGA
 GAATGTTAACTATCATTCAATATTAAATGTAAGGAAATCTGAGAGTGACAAGCTTAGC
 TTTAAATCTGGTATCCAAACTCATTGAGTTTTTTTTTTTTGAGAC
 [A, G]
 AGGTGTCGCTTGTCCCCCAGGCTGGAGTGTAGTGGTGTGATCTGGCTACTGCAACCT
 CCACCTCCCAGGTTCAAGTGGATTCTCTGCCCTCAGCTCTGAAGTTGCTGGGATTACAGG
 CTGCCACCACGCCAGCTAATTGTTGATTTAGTAAAGACGGAGTTCACCTTAT
 TGGCCAGGCTGGCTCAAACCTCTGATCTGTGATCTCCGCCCTGGCTCCCAAAGTG
 CTGGGATTACAGGTGTGAGGCCACTGTTCCGCCCTAATTGAGTTAAAGTGAGGTT
 11670 TGTCATGAACCTCAGATTAAGAATCTGATTAGGAGATTGAAAACAGAAAAGAGAAAT
 GTTAACTATCATTCAATATTAAATGTAAGGAAATCTGAGAGTGACAAGCTTAGCTT
 AATCTGTTATCCAAACTCATTGAGTTTTTTTTTTTTTTGAGACAAGG
 TGTCGCTTGTCCCCCAGGCTGGAGTGTAGTGGTGTGATCTGGCTACTGCAACCTCCA
 CCTCCCAGGTTCAAGTGGATTCTCTGCCCTGAAGTTGCTGGGATTACAGGCTG
 [C, T]
 GCCACCCACGCCAGCTAATTGTTGATTTAGTAAAGACGGAGTTCACCTATTGGC
 CAGGCTGGCTCAAACCTCTGATCTGTGATCTCCGCCCTGGCTCCCAAAGTGCTGG
 GATTACAGGTGTGAGGCCACTGTTCCGCCCTAATTGAGTTAAAGTGAGGTTAAG
 ATGTTAGTCTTAAAGTGAGGTTAGTGAATTTATAAAATGTCATAGCTAAATT
 AAAAGGCCATTGAAACAATTGTAATATAATGTTGAGTTAAAGTGAGGTT
 11688 TAAGAATACTGATTAGGAGATTGAAAACAGAAAAGAGAAATGTTAACTATCATT
 TATTAAATGTAAGGAGTGTGACAAGCTAGCTTAAATCTGGTATCCAAACT

CATTGAGTTTTTTTTTTTTTTTTGAGACAAGGTGTCGTTGCCCCAG
 GCTGGAGTGTAGTGGTGTGATCTGGCTCACTGCAACCTCCACCTCCAGGTCAAGTGA
 TTCTCCTGCCTCAGCCTCTGAAGTTGCTGGATTACAGGCTGCCACAGCCCAGCTA
 [A, G]
 TTTTTGTATTATAGTAAAGACGGAGTTCACCTATTGGCCAGGCTGGTCTCAAAC
 CTGATCTGTGATCCTCCCGCCTGGCCTCCAAAGTGCTGGATTACAGGTTGAGGCA
 CTGTTCCGGCCTAATTGAGTTAAATGTGAGTTAAGATGTTAGCTTAAAGTGG
 GTTAGATGAAATTATAAAAATAGTCAAATAGCTAAATTATAAAAGGCCATTGAAACA
 ATTTGTCAAATATAATGTGATAATTATGTAGTGTCTTATGTAGATTGGTGGTTA

 14938 CATGGTAGTGTGCACCTGTAGTCCAACCACTTGGGAGGCTGAGGTGGGAGGATTGCC
 AGGCCAGGAGTTGAGACCTGGCAGCATATGAAAGACCCCTGCTCTAAAAAACTAAAAA
 AAAAATAGCCAGGTGTTGGTGTGCTTGTGGTCCCAGCTACTCAAGAGGCTGAGGCA
 AGAGGGTTGCTTGAGCCAGAAGTGGAGGCTGCCGTGAAGTGTGATTGACCAACTGC
 TTCAGCCTGGGTGACATAGCAAGACCCGTCTCTGTGGTGGTGGTGGGGTGGGG
 [A, C]
 AGGGATTTAAGAAGGGTTGTGAGGTATGTATTATTATAATGGGCTTTAACATTAC
 CTTCACATCTGGGTGAAATTAAATTGTATCCATTCTCAGTTTCTGTCTGTATATA
 TTAAACTTGGAGACTTAGAGGTCATGGATGTTCTATGAAAAGCAAATGAAGCAGAG
 GGCTGCCCTCTTGTCTAGAGGGCACACTTGTGCAAGAGCATGTTACTGTTATGCA
 TTGCTAGGTTGGAGTTGACTGTATGATCATAGTACTACAACATTAGTTGGCA

 22261 CACCCACAGATAGCTATGTCAAACGTAAGGGTGGAGAAACAGACCCCCAACTTC
 GGGTAGAAAATATGAGGTTATAGTAGATTAGAATCAGAACTACAAAAGCTAGAGGA
 AGTTCTGAACTTGGGTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTGCTCTGTT
 CCCAGGCTGGAGTGAATGGCGCAGTCTCCTCACTGCAACCTCCGCCCTGGGCCA
 [G, A]
 GTGATTCTCCTGCCCTAGGCCACCAAGTAGCTGGGATTACAGGCATCTGCCACCATG
 AGCTAATTGTTGTTAGAGACGGGTTTACCATGTTGCTCAGGCTG
 GTCTGAACCTCTGACCTCAGGTAAATCCACCCACCTGGCCTCCAAAGTGTGGGATTA
 CAGGTGTGAGCCACCGCGCCAGCAAATTGTTCTAAACTAGCGTAATTAG
 TTTTTTCACTTAAGTCAAATTATATTGTAGGATAAAACTAGTGTGATCCAAATT

 22852 ATCCAAATTGAGGAATGAAAGATAAACATTTAAAGTCTTACCATTTGCTAAATT
 GTCTGGCTTTGTACCAAAATTCTGCTTGTGCTGTAAATTATATTGTATATT
 TTCTATCACATTAACTGTGTTGTTGTTGAAATTATAAAACGTTTAAAGCAAAC
 TCAGAACATGAATTCTCACGAATTACAGTATATTACAGTTGAGAAATAACTCTC
 TGTAGTAGTAATTAAATGCTTCAATGCAAGTTAACGTGTCAGTGTGATCACGCTATT
 [G, A]
 GTGTGTCTTGATAAGGGAGGTGGGAAGTTGTGGGTTGATTTATTTGCC
 TCATGTACTGTTGTCATGTTAGTAAACAAATGGTTGCGAGAGAACCGAGTAGCT
 CAAAGATGCTTATACAGAGCACTCAATTCTCATATTATTATAATGGCTTAAATT
 AGCCTTAAATTATAGAAACTCAAATAATTGTTTATTTGAGATGGAGTT
 TGCCCTTATGTCCAGGCTGAAGTACAATGATGTCAGTGTGACTCACTGCAACCTCC

 27253 GCTTAAGCCATGCATGGGTTTATAGGAGATGTAGTCTTACAGTGAGTTGTTATTG
 GCTGTGTTTGTGTTGTATAGCTTACAGCAATGCAGTGCTTTATTAAACATCATT
 TTCTTTCTTTGCACTGATTTTACAGTTCAAGTACTCTGATTGGCACTCAGGG
 TGGAAAGTCTGCTTCTCTTCTAGGTTGCAAGTGAATTGAAATTGCTTTACAAT
 TAATGTCACATTAAATGCTATGTATGTTCTAGGTAGATAAAATTAAACAGTT
 [A, C]
 GAATAAGTTAATTCTCCAGAAATTATATATTAAAGACTCCAAATATAACATCCCCAGTG
 GTATCTGGACTGTTAAAGAAAATTGTTGCTCTAAAGAAATTCAAGTGAAGTCT
 GGTTATAAAAGTCAGAATGCTAATACTTTGGTCAGAGTCACACAGCACTTCAAA
 GCAGCAAGTTAAGGGTAGTGGTGGCCTGTTGAAAGGCACTTGATGAAAATAATC
 TTTAAATTAACTTAGAGATAAAAGAAAAGCAGAGCCAGGTGACCCAGTGGATCA

 28098 CTTAAATTAGCATGTTCTGGCCAGGTGGCTGCGGTGGCTCACGCC
 TTGGGAGGCCAGACGGCGGATCACAGGTCAGAGATTGAGACCATCTGGCTAAC
 GGTGAAACCCGCTCTACTAAACAAATCAGCTGGGTGTTGCCACAGCCT
 GTAGTCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGCTGAAACCCAGGAGG
 TTGCACTGAGCTGAGATGGTGCCTGCACTCCAGCCTGGCAACAGAGCAAGACTGT
 C

FIGURE 3, page 18 of 21

[-, A]

AAAAAAAAAGAAAAAAATAAAAACAAATTAGCATGTTCCCTCTAGAGATCATTGT
 TTCTCAGAGCATGGACCAAGACTCCTGGGGTACCAAGACCCCTCAGGTAGCCATG
 AGGTCAAAATATCCTAATAACTAAGATGTTAGTATTGTAAGGAAATATTACTTGGT
 AATAATACTAATATAAAAGATGTTGCGTTTCAGTGTGACATTGGCTCTGGTACAAA
 AGCATGTGGGTAATTGCTGCTGGTACACATCAAGGCAGCGTAAGCTCCAAAT

28597 GATGTTGCGTTTCAGTGTGACATTGGCTCTGGTACAAAAGCATGTGGTAAATTG
 CTGCTGGCTGGTACACATCAAGGCAGCGTAAGCTCCAAATTGTACTCATGGTGTG
 ATTCTTACCTCTGCCCCACAGGACAAAAACAGCCGTGCCATTATTGAAGAT
 TGTCTTGACAAAACAGTAAATGATTAATTGGAAAATGTTGATCCATGAGTATT
 CTGTTAAAATATTGTGAAGAAATGGGAAGTTACACATAAAACAAATGTTTTTTGTT

[G, T]

TTTTTTTTTTTTTTGAGACAGATTCTGGCTGTTGCCAAGGCTAGGTGCAAGTGGC
 GTCTGCTCCCAGGCTCAAGCTTCTCCACTTCAGCCTCCAAAGTGGCTGGACCTCC
 CAAGTGGATGCGCCATCATGCCCTGGCTGATTTTGATTTTTGAGTGTGACAAGGTCTC
 ACTGTGTTGCACAGGCTGGCTCAAACCTCTGAGCTCAAGCGATGCACTGCCCC
 CCCAAAGTGTGCTGGAGAAAGCACTTTACTGCATACTGGCTAGTGTGTTGGTTATTGG

31431 CTGCATTTTTTTTTTTGGTTGAGATGGAGTCTCGCTCTGTCGCCAGGCTGGA
 GTGCACTCGCAATCTCGGCTCACTGCAGCCTCCACCTCATGGGTTCAAGCGATTCTCC
 ATCTGGTCTCCTGACTAGCTAGGTTACAGCGTGTGCCATCACACCCACTAATT
 GTATTTTAGTAGAGACAGGGTTTACCATGTTGCCAGGCTGGTCTGAACCTCTGATC
 TAAAGTGAACCTCCACCTGGCCTCCAAAGTGTGCTGGATTACATATGTGAGCCACTGC

[C, T, G]

CCTGGCCTCTATATACTCTATAGTACCTGATACTTATTAGGCACTCAATTACAAACATAA
 CTTTTTTTTTTTTTTTTGAGACAGAGACATGCCCTGTGCCCTGGGCTGGAGTGC
 AGTGGCACAGTCTCGGCTCACTGCAACCTCACCTCCGGGTTCAAGTGAATTCTCCTTCC
 TCAGCCTCCGGTAGCTGGGATTACAGGCGCCGCCACAGTCAGCTAATTGGT
 ATTGTTAATAGAGATGAGGTTTACCATCTGGCCAGGCTGATCTCAAACCTCTGACCTT

35704 ATGTTGATCATTGGTGTATAAGATTGGGTGTATTGTTGAAACATTCA
 TTTGTTACTTTCTGTGGCTGGAAGGGATCTTATAGGACACTGTCCTTCATCTTGTCTG
 TCTTCATCTTAATAGGAATTCTTCCATGCCCTGAAGGCTCATTGAAACATT
 TTTGTTGTTTTTATTGGAGATACAGTATTGCTCTGTCCTCCAGGCTGGAGTGCA
 GTGGCGCAGTTGAGCTCACTGCAACCTCCGCCCTGGGTTCAAGTGAATTCTCCTGCCT

[C, T]

AGCCTCCCTAATAGCTGGGATTACATGTTGTAACCCATGCCGGACAATT
 TTTGAGATGAGCCTTGCCTTGTGCCCTGGAGTGCCTGGCAATTCTGGCTC
 GCTCAGCCTCCGCCCTCCAGGTCAAGCAGTTCTTGCCTCAGCCTCCCTGAGTAGCTG
 GGATTACAGCGTGCGCCACACCCCTGCTAATTGGTATTAGTAGAGACAGAG
 TTTCACCATGTTGGTAGGCTGGTCTGAACCTCTGACCTCGTGA
 CTGCCGACTCGG

35728 GATTGGGTGTATTGTTGAAACATTCAATTGGTACTTCTGTGGCTGGAA
 GGGATCTTATAGGACACTGTCCTTCATCTTGCTGTCCTTCATCTTAATAGGAATT
 TTTCCATGCCCTGAAGGCCATTTGAACATTGGTTGTTGTTTATTGGT
 GATACAGTATTGCTCTGTCCTCCAGGCTGGAGTGCAGTGGCGCAGTTGAGCTCACTGC
 ACCTCCGCCCTGGGTTCAAGTGAATTCTCTGCCCTCAGCCTCCCTAATAGCTGGATT

[C, T]

ATGTTGTAACCAACATGCCGGACAATTGGGTTGAGATGGAGGCCCTGCTTGT
 GCCCAGGCTGGAGTGCCTGGTGAATCTTGGCTCGCTGCAGCCTCCGCCCTCCAGGTT
 CAAGCAGTTCTTGCCTCAGCCTCTGAGTAGCTGGGATTACAGGCCTGCGCCACCA
 CCCCTGCTAATTGGTATTGGTAGAGACAGAGTTTACCATGTTGGTAGGCTGGT
 CTGGAACCTCTGACCTCGTGA
 CTGCCCTGGGAAAGTGTGAGCAACAGG

36690 AAAAAAAAAAAAGTAACCAAGGTGTGGTGGTCCATGCCCTGAGTCCTAGCTCCCCAG
 GAGACTGAGGTGGAGGAATGTTGAGGCCAGGACTCAAGGCTGAGTGAGGCAAGATT
 GCACCATTCGACCCACAGCTTGGGGACAGAGTGAGAGACCCCTGTC
 CAAAACAAAATAA
 GGCTGGGCCAGTGGCTGCCCCGCTGTGGTTACGCTTATAGTCCTAGCACTTGGG
 AGGCCAAGGTGGCAGATTGCCCTGAGCTCAGGAGGTCTAAGACCAGCCTGAGCAACATGG

[C, T]

GAAACCTCATTTGCAAAACATACAGAAAAAAACAAAAACACAAAACCTCTAGTT
 GCCAGTTATTGGTATTGATTCTCTAGTGATTCTTCTTTCTGAGACAA

FIGURE 3, page 19 of 21

44831 GGAGGGTTAACATTAAGATAGGAAGCAGGAATTGATTGCTCTAAATTAGAAATTA
TATCCCTAAAATTAAACATGAATACTGGGTGGTAATGATAATTGAGGCAATGTATTT
ATTTGGTGACATTTGCATATATGAAGATTTCTGAATAGGACCTTCAAGATCCTAGG
GGGTTTGGTTGGTTTAATTGTGAGGAATAAAAATCTCTGCCACACTGGCATT
AAGGTGACTGAGGTCAAACGTTCTTAGGTTGAAATAGCAGCCAAAACATTCTCA
[C, T]
GCAGGGGCTTGGGATATGGCTGGCAACACATTGGTGTGGGCTCTTAATTAAATG
ATAAAATTAAAGCTAACACACAAGCCAAAATGAATAGGTTTTTAATTTTATTTC
CTAAACAGGCAATTGAAATACATGGTACAAAATAAGTGGTAAGATAATTGAAATGAA
ATGGACAGAATATTCAATTTCATCTATGAAAATTTCACAATAAAAATCATAGTTACT
TTGTATTATAGGCCTGCTGGATCTTCATCCTCACATAAGGCAACTGACAAATTG

FIGURE 3, page 21 of 21